



# Analyse morphologique d'images pour la modélisation d'environnements urbains

Jorge Eduardo Hernández Londoño

## ► To cite this version:

Jorge Eduardo Hernández Londoño. Analyse morphologique d'images pour la modélisation d'environnements urbains. Mathématiques [math]. École Nationale Supérieure des Mines de Paris, 2009. Français. NNT : . pastel-00005974

**HAL Id: pastel-00005974**

**<https://pastel.archives-ouvertes.fr/pastel-00005974>**

Submitted on 4 May 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





ED n° 431 : Information, communication, modélisation et simulation

# **T H È S E**

pour obtenir le grade de  
**Docteur de l'École des Mines de Paris**  
Spécialité «Morphologie Mathématique»

Présentée et soutenue publiquement  
par  
**Jorge Eduardo HERNÁNDEZ LONDOÑO**

le 14 décembre 2009

**Analyse morphologique d'images pour la modélisation  
d'environnements urbains**

Directeurs de thèse : Beatriz MARCOTEGUI ITURMENDI et Dominique JEULIN

## **Jury :**

Philippe SALEMBIER	<i>Rapporteur</i>
Nicolas PAPARODITIS	<i>Rapporteur</i>
Fernand MEYER	<i>Examineur</i>
François GOULETTE	<i>Examineur</i>
Raouf BEN JEMMA	<i>Examineur</i>
Philippe JAROSSAY	<i>Examineur</i>
Beatriz MARCOTEGUI ITURMENDI	<i>Directrice de thèse</i>
Dominique JEULIN	<i>Directeur de thèse</i>



# Table des matières

---

<b>Table de matières</b>	<b>iv</b>
<b>Table de figures</b>	<b>x</b>
<b>Liste de tableaux</b>	<b>1</b>
<b>Remerciements</b>	<b>2</b>
<b>Résumé</b>	<b>4</b>
<b>Abstract</b>	<b>6</b>
<b>Introduction</b>	<b>7</b>
 <b>I Modélisation Urbaine</b>	 <b>10</b>
 <b>1 Projet TerraNumérica</b>	 <b>11</b>
1.1 Introduction . . . . .	11
1.2 Architecture haussmannienne . . . . .	12
1.3 Données . . . . .	13
 <b>2 Etat de l'Art : Modélisation d'Environnements Urbains</b>	 <b>17</b>
2.1 Introduction . . . . .	17
2.2 Techniques d'acquisition . . . . .	18
2.3 Techniques de recalage et de filtrage . . . . .	23
2.4 Techniques d'infographie . . . . .	25

2.5	Techniques de reconstruction . . . . .	27
2.6	Discussion . . . . .	31
<b>II</b>	<b>Extensions de l'Ouverture Ultime</b>	<b>32</b>
<b>3</b>	<b>Ouverture Ultime par Attributs</b>	<b>33</b>
3.1	Introduction . . . . .	33
3.2	Opérateurs connexes et zones plates . . . . .	34
3.3	Ouverture par attributs . . . . .	34
3.4	Ouverture ultime . . . . .	35
3.5	<i>Max-Tree</i> . . . . .	37
3.6	Ouverture ultime sur des images réelles . . . . .	40
3.7	Problèmes de l'ouverture ultime . . . . .	40
3.8	Discussion . . . . .	46
<b>4</b>	<b>Ouverture Ultime avec l'Information de Forme</b>	<b>47</b>
4.1	Introduction . . . . .	47
4.2	Similarité entre formes . . . . .	48
4.3	Ouverture ultime avec l'information de forme . . . . .	49
4.4	Exemple d'une image synthétique . . . . .	51
4.5	Application : Images de façade . . . . .	53
4.6	Discussion . . . . .	58
<b>5</b>	<b>Ouverture Ultime Hiérarchique</b>	<b>60</b>
5.1	Introduction . . . . .	60
5.2	Hiérarchisation de l'ouverture ultime . . . . .	61
5.3	Application : Images de façade . . . . .	68
5.4	Discussion . . . . .	70
<b>6</b>	<b>Ouverture Ultime par <math>\Delta</math>-attributs</b>	<b>73</b>
6.1	Introduction . . . . .	73
6.2	Ouverture Ultime par $\Delta$ -attribut . . . . .	74

6.3	Application : Images de texte enfoui . . . . .	80
6.4	Discussion . . . . .	85
<b>7</b>	<b>Applications</b>	<b>86</b>
7.1	Introduction . . . . .	87
7.2	Segmentation d'images de façade . . . . .	87
7.3	Détection de texte enfoui . . . . .	98
7.4	Segmentation d'images de cellules . . . . .	107
7.5	Discussion Générale . . . . .	109
<b>III</b>	<b>Analyse des Images Couleurs</b>	<b>113</b>
<b>8</b>	<b>Segmentation des Façades par Bâtiment</b>	<b>114</b>
8.1	Introduction . . . . .	114
8.2	Séparation par bâtiments . . . . .	116
8.3	Détection du ciel . . . . .	118
8.4	Séparation du rez-de-chaussée et de la chaussée . . . . .	122
8.5	Discussion . . . . .	123
<b>9</b>	<b>Extraction de la Sémantique d'une Façade</b>	<b>126</b>
9.1	Introduction . . . . .	126
9.2	Division de la façade . . . . .	126
9.3	Détection de fenêtres et de balcons . . . . .	129
9.4	Modélisation procédurale . . . . .	134
9.5	Discussion . . . . .	135
<b>IV</b>	<b>Analyse des Nuages de Points</b>	<b>137</b>
<b>10</b>	<b>Segmentation de Nuages de Points</b>	<b>138</b>
10.1	Introduction . . . . .	138
10.2	Projection sur images . . . . .	139
10.3	Segmentation par îlots . . . . .	141

10.4	Segmentation façades-sol . . . . .	143
10.5	Segmentation de façades en bâtiments . . . . .	148
10.6	Discussion . . . . .	149
<b>11</b>	<b>Détection et Classification d'Artefacts</b>	<b>152</b>
11.1	Introduction . . . . .	152
11.2	Détection et filtrage . . . . .	153
11.3	Classification d'artefacts . . . . .	156
11.4	Discussion . . . . .	159
<b>12</b>	<b>Segmentation de la Chaussée</b>	<b>163</b>
12.1	Introduction . . . . .	163
12.2	Projection sur images sans artefacts . . . . .	164
12.3	Segmentation des points du sol . . . . .	165
12.4	Sélection de frontières des régions . . . . .	169
12.5	Discussion . . . . .	170
<b>13</b>	<b>Conclusions et Perspectives</b>	<b>172</b>
13.1	Apports de cette thèse . . . . .	172
13.2	Perspectives . . . . .	174
	<b>Bibliographie</b>	<b>176</b>

# Table des figures

1.1	Screenshot Google Maps. En rouge la zone de test . . . . .	13
1.2	Données ©IGN – Rue Soufflot . . . . .	14
1.3	Séquences d’images Données ©IGN – Rue Soufflot . . . . .	14
1.4	Données ©IGN . . . . .	15
1.5	Données ©CAOR – Rue Soufflot . . . . .	15
1.6	Données ©MENSI – Rue Soufflot Bâtiment N 23 . . . . .	16
2.1	Les étapes de la modélisation d’environnements urbains. . . . .	19
3.1	(a) Les mesures du plus petit rectangle (boîte englobante) qui contient $C$ telles que : la surface, la diagonale, la hauteur (le diamètre de Ferret vertical), la largeur (le diamètre de Ferret horizontal), l’élongation, etc. (b) Les mesures de la plus petite circonférence qui enferme $C$ telles que : la surface, le rayon, le diamètre (le diamètre maximum de Ferret), etc. (c) Les mesures du rectangle le plus grand contenu dans $C$ . (d) Les mesures de la circonférence la plus grande contenue dans $C$ . . . . .	36
3.2	(a) Image Synthétique, (b) $\mathcal{J}_h(I) \forall h \in E$ , (c) <i>Max-Tree</i> . . . . .	37
3.3	(a) <i>Max-Tree</i> et trait pointillé aux branches à élaguer (b) ouverture par hauteur $\gamma_T(I)$ , où $T : \kappa_{C_h^k} \geq \lambda$ . . . . .	38
3.4	(a) <i>Max-Tree</i> et résidu $r_\lambda(C_h^k) \forall h \in E$ , (b) les images intermédiaires de $r_\lambda(I)$ , (c) <i>Max-Tree</i> et $R_\theta(C_h^k)$ , $q_\theta(C_h^k)$ et les images résultats de : (d) $R_\theta(I)$ , et (e) $q_\theta(I)$ de l’ouverture ultime par hauteur. . . . .	39
3.5	Ouverture ultime sur des images réelles. . . . .	41
3.6	Exemple : Problème dû aux structures imbriquées (a) Image Synthétique, (b) Représentation en relief d’ $I$ , (c) Séries de transformées et (d) Séries d’indicatrices. . . . .	42
3.7	Exemple : Problème dû aux transitions graduelles (a) Image Synthétique, (b) Représentation en relief d’ $I$ , (c) Séries de transformées et (d) Séries d’indicatrices. . . . .	43
3.8	Exemple : Problème dû aux problèmes de fuites (a) Image Synthétique, (b) Représentation en relief d’ $I$ , (c) Séries de transformées et (d) Séries d’indicatrices. . . . .	43

3.9	Problèmes de l'ouverture ultime sur des images réelles. . . . .	45
4.1	Segmentation de l'image synthétique avec $\theta^\Omega$ . Eq. 4.7 comme fonction du facteur de forme. . . . .	52
4.2	Segmentation de l'image synthétique avec $\theta^\Omega$ . Eq. 4.8 comme fonction du facteur de forme. . . . .	53
4.3	Segmentation de l'image synthétique avec $\theta^\Omega$ . Eq. 4.9 comme fonction du facteur de forme. . . . .	53
4.4	(a) Image Originale, (b)-(c) $R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime, (d) Image Mosaïque sur $q_\theta(I)$ . (e)-(g) $F_\theta^\Omega(I)$ , $R_\theta^\Omega(I)$ et $q_\theta^\Omega(I)$ issues de l'ouverture ultime avec l'information de forme, (h) Image Mosaïque sur $q_\theta^\Omega(I)$ . ‡ Correction gamma de 2. L'image mosaïque (§) est la valeur moyenne de l'image originale sur la partition générée par la ligne de partage des eaux du gradient en utilisant comme marqueur $q_\theta$ . . . . .	55
4.5	(a) Image Originale, (b)-(c) $R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime, (d) Image Mosaïque sur $q_\theta(I)$ . (e)-(g) $F_\theta^\Omega(I)$ , $R_\theta^\Omega(I)$ et $q_\theta^\Omega(I)$ issues de l'ouverture ultime avec l'information de forme, (h) Image Mosaïque sur $q_\theta^\Omega(I)$ . ‡ Correction gamma de 2. . . . .	56
4.6	(a) Image Originale, (b)-(c) $R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime, (d) Image Mosaïque sur $q_\theta(I)$ . (e)-(g) $F_\theta^\Omega(I)$ , $R_\theta^\Omega(I)$ et $q_\theta^\Omega(I)$ issues de l'ouverture ultime avec l'information de forme, (h) Image Mosaïque sur $q_\theta^\Omega(I)$ . ‡ Correction gamma de 2. . . . .	57
4.7	(a-b) Images de l'exemple de la Figure 4.5. (c-d) Images de l'exemple de la Figure 4.6. . . . .	59
5.1	(a) Image $I$ , (b) Masque $X$ . (c) Ouverture ultime géodésique de $I$ en $X$ . (d) Ouverture ultime du masquage de $I$ en $X$ . (e) Ouverture ultime de $I$ masquée par $X$ . ‡ Correction gamma de 3 . . . . .	62
5.2	(a) Image de façade, (b) indicatrice $q_\theta$ de l'opérateur classique $\theta(I)$ et niveau zéro d'hierarchie, (c) $q_{\theta_T^3}$ indicatrice de l'opérateur hiérarchique $\theta_T^3(I)$ , (d)-(f) images d'étiquettes $V^k$ qui vérifient le critère et (g)-(i) indicatrice $q_{\theta_{V^k}}$ de l'opérateur géodésique $\theta_{V^k}(I)$ . . . . .	65
5.3	(a) Image Synthétique, (b) Max-Tree et tableau d'attributs . . . . .	67
5.4	(a) $R_{\theta_T}(C_h^k)$ , $q_{\theta_T}(C_h^k)$ et les images résultats de : (b) $R_{\theta_T}(I)$ , et (c) $q_{\theta_T}(I)$ de l'ouverture ultime hiérarchique par hauteur. $T$ : Surface $(C_h^k) > 2$ . . . . .	67
5.5	(a) Image Originale, (b)-(c) $R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime, (d) Image Mosaïque sur $q_\theta(I)$ . ‡ Correction gamma de 2. . . . .	68
5.6	(a)-(b) $R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime, (c) Image Mosaïque sur $q_\theta(I)$ . Correction gamma de 2. . . . .	70
5.7	Résultats de l'opérateur $\theta_T$ . . . . .	71
6.1	Exemple d'un profil avec des transitions graduelles. . . . .	74
6.2	Calcul de l'ouverture ultime par $\Delta$ —attribut depuis l'ouverture la plus petite jusqu'à la plus grande. . . . .	76
6.3	Calcul de l'ouverture ultime par $\Delta$ —attribut depuis l'ouverture la plus grande jusqu'à la plus petite. . . . .	77



6.4	Exemple : Transitions graduelles (a) Image Synthétique, (b) et (c) Ouverture ultime par $\Delta$ -attribut. . . . .	79
6.5	(a) <i>Max-Tree</i> d'un profil avec des transitions graduelles. (b) Calcul de résidus classique. (c)-(d) Calcul de l'accumulation des résidus avec $\Delta = 1$ . . . . .	80
6.6	(a) Image Originale. (b)-(c) Ouverture ultime classique. (d)-(e) Ouverture ultime avec $\Delta$ -attribut $\Delta = 1$ . ‡ Correction gamma de 2. . . . .	81
6.7	Exemple : Transitions graduelles (a) Image Synthétique, (b) et (c) Ouverture ultime par $\Delta$ -attribut. ‡ Correction gamma de 2. . . . .	82
6.8	Série de résidus de deux pixels. . . . .	83
6.9	Problème par des structures imbriquées. (a) Image Originale. (b)-(c) Ouverture ultime classique. (d)-(e) Ouverture ultime avec $\Delta$ -attribut $\Delta = 1$ . ‡ Correction gamma de 2. . . . .	84
6.10	Exemple : Filtrage des valeurs perchées (a) Image Originale, Ouverture ultime par $\Delta$ -attribut par : (b) inondation et (c) <i>Max-Tree</i> . . . . .	84
7.1	Résultats des mesures de qualité d'une segmentation. . . . .	90
7.2	Résultats des mesures normalisées de qualité d'une segmentation. . . . .	91
7.3	(a) Image Originale, (b) Image Gradient Inversée, (c)-(d) $R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime, (e) Image Mosaïque sur $q_\theta(I)$ . ‡ Correction gamma de 2. § L'image mosaïque est la valeur moyenne de l'image originale sur la partition générée par la ligne de partage des eaux du gradient en utilisant comme marqueur $q_\theta$ . . . . .	92
7.4	(a)-(d) $F_\theta^\Omega(I)$ , $R_\theta^\Omega(I)$ et $q_\theta^\Omega(I)$ issues de l'ouverture ultime avec l'information de forme, (e) Image Mosaïque sur $q_\theta^\Omega(I)$ . ‡ Correction gamma de 2. . . . .	93
7.5	Critère de surface. Mesures de qualité moyenne de l'ouverture ultime hiérarchique sur la base de test de 50 images. . . . .	94
7.6	Critère d'hétérogénéité couleur. Mesures de qualité moyenne de l'ouverture ultime hiérarchique sur la base de test de 50 images. . . . .	94
7.7	$R_\theta(I)$ et $q_\theta(I)$ issues de l'ouverture ultime hiérarchique et image Mosaïque sur $q_\theta$ . ‡ Correction gamma de 2. . . . .	95
7.8	Mesures de qualité moyenne de l'ouverture ultime par $\Delta$ - attributs sur la base de test de 50 images. . . . .	96
7.9	$R_\theta^\Delta(I)$ et $q_\theta^\Delta(I)$ issues de l'ouverture ultime hiérarchique et image Mosaïque sur $q_\theta^\Delta$ . ‡ Correction gamma de 2. . . . .	97
7.10	Mesures de qualité moyenne sur la base de test de 50 images. . . . .	97
7.11	Exemple d'image de texte dans les deux polarités. . . . .	98
7.12	Exemple de l'ouverture ultime pour la détection de texte enfoui sur les deux polarités de l'image. . . . .	99
7.13	Histogrammes de 5000 caractères analysés. . . . .	100
7.14	Images de test pour la détection de texte enfoui. . . . .	101
7.15	Résultats des ouvertures ultimes sur la Figure 7.14(a). ‡ Correction gamma de 2. . . . .	104

7.16	Résultats des ouvertures ultimes sur la Figure 7.14(b). ‡ Correction gamma de 2. . . . .	105
7.17	Résultats des ouvertures ultimes sur la Figure 7.14(c). ‡ Correction gamma de 2. . . . .	106
7.18	Résultats des ouvertures ultimes sur une image de cellules. ‡ Correction gamma de 2. . . . .	109
7.19	Evolution de différents seuils de manière décroissante. . . . .	109
7.20	Analyse de l'ouverture ultime sur l'image de luminance et l'image gradient. . . . .	110
7.21	Effet du paramètre $\alpha$ , fonction de forme de l'équation 7.4. $q_\theta(I, I^c)$ des ouvertures ultimes avec l'information de forme. . . . .	111
8.1	Image de façade. . . . .	115
8.2	Diagramme de la segmentation de façades par bâtiment. . . . .	115
8.3	(a) Gradient couleur vertical et sa projection. (b) Gradient filtré et sa projection. (c) Projection verticale $P_v(\widetilde{G}_v)$ sur l'image de départ. (d) Division verticale de la façade. . . . .	116
8.4	Sélection des divisions par <i>k-means</i> . . . . .	117
8.5	Fausse détection inter-façade. . . . .	118
8.6	Diagramme de la détection du ciel. . . . .	120
8.7	Résultats de détection du ciel. . . . .	120
8.8	Cas d'échec de la détection du ciel. . . . .	121
8.9	Détection et lignes d'élimination du ciel. . . . .	122
8.10	(a) Gradient couleur horizontal et (b) gradient filtré. (c) Projection horizontale de $\widetilde{G}_h$ et division horizontale de façade, (d) Projection horizontale de $\widetilde{G}_v$ et division horizontale raffinée de façade. . . . .	123
8.11	Analyse horizontale par étages. . . . .	124
8.12	Ligne inter-façade occlue pour un arbre. . . . .	124
8.13	Résultats de la segmentation de façades par bâtiments. . . . .	125
9.1	Diagramme simplifié de subdivision d'une façade haussmannienne. . . . .	127
9.2	(a) Projection horizontale de $\widetilde{G}_v$ et division horizontale raffinée de façade. (b) Détection de la dalle. . . . .	128
9.3	Effets du filtrage vertical sur le profil. . . . .	129
9.4	Division de la façade en étages + dalle et en travées. . . . .	130
9.5	Diagramme simplifié de subdivision d'une travée. . . . .	130
9.6	Localisation de la bordure verticale des fenêtres. . . . .	131
9.7	Localisation de fenêtres. . . . .	132
9.8	Adéquation de l'image de luminance pour renforcer la couleur sombre des balcons. . . . .	133
9.9	(a) Projection horizontale et verticale de l'image. (b) Détection horizontale de balcons et (c) Détection verticale de balcons. . . . .	133

9.10	Instanciation du modèle de la façade. . . . .	134
9.11	Rendu 3D du modèle de la façade. . . . .	134
9.12	Détection de dalles en compensant la profondeur des balcons. . . . .	135
9.13	Robustesse aux obstacles. . . . .	136
10.1	(a) Nuages de points - vue 3D et vue supérieure. Projection sur images (code couleurs au rang) : (b) la valeur maximale – rang (0, 23651mm), (c) la différence max – min – rang (0, 18816mm) et (d) l'accumulation – rang (0, 121fois). Données ©IGN Tronçon de la Rue Soufflot côté pair. . . . .	140
10.2	Les étapes de la segmentation par îlots. . . . .	141
10.3	Alignement global de façades. (a)- (c) rue Soufflot, (b)- (d) place du Panthéon. . . . .	142
10.4	Images intermédiaires de la segmentation par îlots de rues alignées. . . . .	142
10.5	Segmentation par îlots de rues non-alignées. Données ©CAOR Fin de la Rue Soufflot – début de la Place du Panthéon. . . . .	143
10.6	Segmentation par îlots de rues non-alignées. Données ©IGN Place du Panthéon. . . . .	144
10.7	Illustration de l'acquisition . . . . .	145
10.8	Assignment des pixels manquants avec la valeur minima. . . . .	145
10.9	Assignment des pixels manquants avec le remplissage de trous. . . . .	146
10.10	Estimation des pixels manquants en utilisant le remplissage de trous. . . . .	147
10.11	Segmentation Façade-Sol. . . . .	148
10.12	Les étapes de la segmentation de façades en bâtiments. . . . .	149
10.13	Segmentation de façades en bâtiments. Données 3D ©IGN, Image 2D ©MENSI Rue Soufflot - Bâtiment N-13. . . . .	150
10.14	Segmentation des étages de façades. . . . .	150
11.1	Détection d'artefacts du profil 10.9. . . . .	154
11.2	(a) Illustration d'un artefact proche du capteur. (b) Problème d'inversion de points à zero. (c) Problème des artefacts collés au bord . (d) Résultat final. . . . .	155
11.3	Les étapes de la détection d'artefacts. . . . .	155
11.4	Exemples d'artefacts connectés. . . . .	156
11.5	(a) Détection initiale et (b) Séparation d'artefacts. . . . .	157
11.6	Sélection de caractéristiques. . . . .	158
11.7	Résultats des méthodes de détection et classification d'artefacts. . . . .	160
11.8	Points sol assignés comme artefact. . . . .	161
11.9	Illustration d'artefacts cachés. . . . .	162

11.10	Illustration d'une séparation indésirable. . . . .	162
12.1	Modélisation : <b>(a)</b> d'un bâtiment, <b>(b)</b> avec des lampadaires et <b>(c)</b> avec la chaussée. . . . .	164
12.2	<b>(a)</b> Points du sol manqués en vert et <b>(b)</b> Points qui appartiennent au plan estimé en cyan. . . . .	165
12.3	Projection sur images. . . . .	166
12.4	Diagramme de la segmentation de la chaussée. . . . .	167
12.5	Segmentation de $\lambda$ -zones plates. . . . .	168
12.6	Illustration de l'information manquante dans la séparation trottoir-rue. . . . .	168
12.7	Remplissage directionnel des trous en utilisant la ligne de la façade. <b>(a)</b> La ligne de la façade en vert. . . . .	169
12.8	Le bord local de l'épaisseur $d = 1$ . . . . .	169
12.9	Diagramme de la segmentation finale de la chaussée. . . . .	170
12.10	Projection de la segmentation sur le nuage de points original . . . . .	170

# Liste des tableaux

---

7.1	Mesures d'évaluation de résultats de segmentation d'images. . . . .	88
8.1	Matrice de Confusion de la détection du ciel. . . . .	121
11.1	Matrice de Confusion ALD. . . . .	158
11.2	Matrice de Confusion SVM. . . . .	159

## Remerciements

---

Tout d'abord je souhaite adresser un très grand merci à M. Fernand Meyer, directeur du Centre de Morphologie Mathématique et à M. Dominique Jeulin, directeur de l'école doctorale pour m'avoir accueilli dans leur laboratoire et pour m'avoir prodigué de précieux conseils au cours de ces trois années.

Je souhaite également exprimer ma plus sincère gratitude à Mme. Beatriz Marcotegui pour m'avoir encadré et donné son avis toujours très pertinent, d'avoir partagé ses connaissances et son inépuisable énergie avec moi. Merci de m'avoir permis de réaliser mon travail dans des conditions idéales et merci pour toutes ces heures de discussion scientifique qui m'ont été ô combien profitables. Son soutien constant et l'intérêt qu'elle a porté à mon travail m'ont énormément aidé dans mon apprentissage de la recherche. Je la remercie également de la confiance qu'elle m'a témoignée en acceptant de suivre mes initiatives et j'espère en avoir été digne.

Je tiens à exprimer mes remerciements à chacun des membres du jury, qui ont accepté d'évaluer mon travail de thèse. J'exprime ma gratitude à M. Nicolas Paparoditis et à M. Philippe Salembier, pour l'intérêt qu'ils ont porté à mon travail en tant que rapporteurs de thèse, ainsi qu'à Mr Raouf Ben Jemma et à Mr Philippe Jarossay pour avoir accepté de participer à ce jury de thèse, de même qu'à Mr François Goulette sa participation et pour les discussions que nous avons eues pendant le projet.

Je profite également de l'occasion pour remercier chaleureusement Mme Catherine Moysan, qui, grâce à sa bonne humeur et à son enthousiasme, m'a permis d'effectuer cette thèse dans de très bonnes et très agréables conditions de travail. Merci pour sa relecture méticuleuse du manuscrit en améliorant la qualité des différentes sections.

Plus généralement, j'adresse mes sincères remerciements à toutes les personnes du CMM, trop nombreuses pour être citées, pour les bons moments passés ensemble dans le travail comme en dehors. Je garderai notamment toujours la saveur de ces pauses cafés, espaces de vraie discussion. Merci à Charles, Matthieu, Romaric, Louise, Serge,... à leur enthousiasme, parfois débordant, qui contribuent à l'ambiance de travail agréable et propice à la recherche. Je me souviendrai également de tous ceux qui ont jeté un coup d'œil à l'orthographe du manuscrit. Je

---

remercierai plus particulièrement Edwige pour son aide constante et minutieuse tout au long de la rédaction de la thèse aux heures du tandem.

Finalement, j'exprime mon immense gratitude à ma famille, et notamment à mes parents, pour m'avoir soutenu durant toutes ces années. Sans leur aide, ce projet n'aurait pas pu aboutir. Je leur serai toujours redevable de tous les efforts qu'ils ont fournis à mon égard. Enfin, si je n'avais qu'une personne à remercier, ce serait celle qui partage ma vie depuis quelques années et qui, au cours de cette thèse, est devenue mon épouse. Alors mille mercis à Carol pour tout le bonheur qu'elle m'apporte, pour sa patience, son soutien. . . Elle mérite amplement que je lui dédie ce travail.

## Résumé

---

La modélisation d'environnements urbains permet le développement de nombreuses applications telles que : la navigation 3D, la planification des projets d'aménagement urbain et la création de scénarii pour l'industrie du cinéma ou des jeux vidéo. L'enjeu technologique est de réduire le temps nécessaire pour une modélisation réaliste et fidèle à la réalité. Cette thèse s'inscrit dans le cadre du projet TerraNumérica, du pôle de compétitivité mondial Cap Digital, dont l'objectif est de développer une plateforme de production et d'exploitation d'environnements synthétiques urbains. Dans ce contexte, nous sommes confrontés à des problèmes de segmentation de scènes urbaines.

Dans une première partie, nous avons montré l'intérêt de l'opérateur d'ouverture ultime en tant qu'opérateur de segmentation générique. Ensuite nous avons proposé plusieurs améliorations qui le rendent plus robuste à des problèmes de masquage et d'images floues. L'efficacité de ces améliorations est démontrée non seulement dans le contexte de l'étude mais également dans le cadre de deux autres applications : la localisation de texte enfoui et la segmentation de cellules.

La deuxième partie est consacrée à l'analyse de façades. La modélisation de celles-ci est réalisée à l'échelle du bâtiment. Néanmoins, lors de l'acquisition, plusieurs bâtiments apparaissent dans une même image. Nous proposons une méthode capable de séparer de manière automatique les différentes façades contenues dans l'image. Ensuite, nous nous intéressons à l'extraction d'information sémantique de la façade. Elle est ainsi segmentée en étages, travées, fenêtres... afin d'alimenter un modèle de bâtiment fidèle à la réalité.

Par ailleurs, nous disposons également de nuages de points issus de systèmes mobiles. La troisième partie de cette thèse est consacrée à l'analyse des données tridimensionnelles. Les systèmes mobiles fournissent un flux continu de données. Nous proposons une méthode automatique pour segmenter ce flux en îlots, évitant d'une part les éventuels problèmes de mémoire et d'autre part permettant l'analyse de données cohérentes, car l'îlot constitue une entité sémantique autour de laquelle se trouve le trottoir. Pour une modélisation de la scène au niveau du sol, il est nécessaire d'introduire certains éléments du mobilier urbain (lampadaires, panneaux de signalisation) ainsi que le trottoir. Nous présentons des outils pour la détection et la classification d'artefacts (tout objet autre



---

que les façades et le sol). Ces outils permettent le filtrage de données pour faciliter l'étape de modélisation et la réintroduction de certains éléments, améliorant le réalisme visuel de la scène urbaine modélisée. Nous proposons également une méthode automatique pour la segmentation du trottoir.

# Abstract

---

Urban modeling allows the development of many applications such as : 3D navigation, urban planning, scene modeling for the entertainment industry , etc. The technological challenge is to increase the productivity and the realism of urban modeling. This thesis is developed in the framework of Cap Digital Business Cluster TerraNumerica project. The project aims at developing a production and exploitation platform, by allowing the definition and visualization of synthetic urban scenes. In this context, our main contribution to the project consists in segmenting color images and point clouds in order to assist urban modeling.

This document is mainly divided into three parts. In the first part, we have shown the ultimate opening benefits as a generic operator of segmentation. Then we have proposed several improvements to make it more robust to masking and leakage problems. The performance of these improvements is illustrated in our study framework, scene-text detection and cell segmentation.

The second part is focused on façade analysis. Façade modeling is performed on the scope of a single building. However, during the image acquisition process, several buildings appear in the same image. We propose an automatic method to separate different façades included in the image. Then, we focus on the semantic extraction from the façade. It consists in segmenting it by floors, windows, balconies... to provide a realistic building model.

The third part of this thesis is focused on point cloud analysis. In urban modeling, it is necessary to introduce some elements such as street furniture and pavement. We present tools for the detection and classification of artifacts. These tools allow : 1- the filtering of data in order to facilitate the modeling process and 2- the re-introduction of some elements (lampposts, sign boards, bus stop, etc), improving visual realism in the modeled scenes. We also propose an automatic method for pavement segmentation.

# Introduction

---

## Contexte

« *Promenons-nous dans des villes virtuelles comme si nous y étions !* » Cette phrase paraît extraite d'un ouvrage de Jules Verne ; cependant, cela n'est plus de la science-fiction. Avec la modélisation 3D du milieu urbain et l'exploration des géo-portails (IGN - Geoportail 3D, Google Earth, Street View, Microsoft Virtual Earth), l'idée d'une visite virtuelle de villes 3D devient de plus en plus réelle. Mais la modélisation d'environnements urbains ne s'arrête pas uniquement à la navigation 3D car on trouve notamment d'autres applications telles que les études environnementales, l'aménagement urbain, la conservation du patrimoine, l'industrie du cinéma et des jeux vidéo, les simulateurs et les applications militaires. Ces modèles urbains sont principalement conçus par des méthodes purement infographiques, dont le temps de réalisation devient coûteux. Par ailleurs, les nouvelles technologies d'acquisition de données se présentent comme des alternatives de perception de notre monde réel. Ainsi, nous sommes confrontés à des enjeux économiques afin de réduire le coût de production de la modélisation et à des enjeux technologiques afin d'intégrer les données réelles dans un monde synthétique. De nombreux projets de recherche au niveau national et mondial, pour augmenter le degré d'automatisation de la modélisation, témoignent de l'intérêt des diverses communautés scientifiques et industrielles dans l'exploitation d'environnements urbains synthétiques. D'ailleurs ce travail est développé dans le cadre du projet TerraNumérica du pôle de compétitivité mondial Cap Digital, dont l'objectif général est de développer une plateforme de production et d'exploitation d'environnements urbains synthétiques.

D'un point de vue général, notre travail se porte sur l'analyse d'images et de données laser afin de produire une lecture sémantique de ce type de données. En effet, deux objectifs sont attendus, tout d'abord la segmentation de ces données facilitant leur interprétation, et l'interprétation de l'information afin d'alimenter une modélisation basée sur des règles.

---

## Sommaire

Après cette introduction décrivant le contexte et le propos de ce travail, le document se divise en quatre grandes parties :

### Première Partie

La première partie décrit le projet TerraNemérica dans lequel s'inscrit cette thèse, ainsi que les données sur lesquelles nous allons travailler. Nous illustrons également l'architecture haussmannienne étudiée dans ce travail. En fin, un état de l'art des grands axes de modélisation urbaine est présenté, permettant de positionner les méthodes proposées vis-à-vis de la littérature.

### Deuxième Partie

Cette partie est dédiée à une analyse approfondie de l'opérateur d'ouverture ultime. L'ouverture ultime est un opérateur morphologique non paramétrique, idéal pour l'extraction des composantes les plus contrastées à différentes échelles. Tout d'abord nous présentons les définitions morphologiques afin d'étudier l'opérateur. Ensuite nous mettons en lumière les problèmes de l'opérateur dus aux structures imbriquées et aux transitions graduelles. Afin de réduire ces problèmes, trois améliorations sont proposées. La première combine l'opérateur avec l'information de forme *a priori*, pour une application donnée. Dans la deuxième nous introduisons l'ouverture ultime hiérarchique, ainsi que l'ouverture ultime géodésique. Finalement, la troisième proposition définit l'ouverture ultime par  $\Delta$ -attributs afin d'accumuler des résidus consécutifs non nuls. Les trois propositions sont validées non seulement dans le contexte de l'étude mais également dans le cadre de deux autres applications : la localisation de texte enfoui et la segmentation de cellules.

### Troisième Partie

Cette partie est consacrée à l'analyse d'images de façade comme phase préalable à la modélisation procédurale. La modélisation de façades est effectuée bâtiment par bâtiment ; cependant, les images prises au sol présentent plusieurs façades de bâtiments. Nous décrivons une méthode de découpage automatique pour séparer les différentes façades de l'image. Une fois que l'image d'une façade est isolée, nous continuons l'analyse de la façade avec l'extraction de ses éléments caractéristiques (sémantique). Ainsi, nous proposons des méthodes pour la division de la façade en étages et en travées, et pour la détection des fenêtres et des balcons. Une dernière section illustre l'intégration de ces méthodes dans une plateforme de modélisation procédurale.

---

## Quatrième Partie

La quatrième partie expose l'analyse de nuages de points issus de systèmes laser (fixes et mobiles). Dans le cas particulier des systèmes mobiles, les nuages de points sont acquis en continu, donnant lieu à de volumes de données importantes. Ainsi, une première étape d'analyse est la séparation de ces points en îlots. Nous avons choisi les îlots comme unité d'analyse pour deux raisons : 1- le volume de données est raisonnable et peut être traité par un ordinateur standard en évitant les éventuels problèmes de mémoire, 2- un îlot correspond à une unité sémantique, autour de laquelle se trouve le trottoir que nous allons localiser par la suite. Par ailleurs, les acquisitions sont réalisées sans intervenir dans la vie normale de la ville, ainsi plusieurs artefacts (tous les objets autres que les façades et le sol) représentent un obstacle durant la modélisation. C'est pourquoi, nous décrivons une méthode pour la détection et la classification d'artefacts. De plus, divers artefacts détectés comme ceux appartenant au mobilier urbain (lampadaires, panneaux de signalisation), peuvent être utilisés pour compléter la modélisation urbaine. Par ailleurs nous proposons une méthode pour la segmentation du trottoir.

Dans le chapitre final, la conclusion de ce manuscrit résume la contribution et discute les perspectives de notre travail du point de vue de la recherche.

**Première partie**

# **Modélisation Urbaine**

# **Projet TerraNumérica**

---

## **Sommaire**

<b>1.1</b>	<b>Introduction</b>	<b>11</b>
<b>1.2</b>	<b>Architecture haussmannienne</b>	<b>12</b>
<b>1.3</b>	<b>Données</b>	<b>13</b>
1.3.1	Matis de l'IGN	13
1.3.2	CAOR Mines-ParisTech	15
1.3.3	MENSI - Trimble	15

## **1.1 Introduction**



Cette thèse s'inscrit dans le cadre du projet TerraNumérica du pôle de compétitivité mondial Cap Digital. Ce projet a pour objectif le développement d'une plateforme de production et d'exploitation d'environnements synthétiques. Il se concentre sur la modélisation du territoire et de ses ressources (zones urbaines, périurbaines, patrimoine culturel, touristique et services urbains) ainsi que sur le développement d'une plateforme permettant la visualisation de ces environnements pour différents supports de consultation (ordinateur personnel, réseau de stations de travail, terminaux mobiles, applications en ligne) sous différentes modalités (réalité virtuelle, réalité augmentée, etc.). Les enjeux technologiques du projet sont : l'automatisation de la numérisation de grandes bases de données urbaines 3D, la structuration et la visualisation optimisée des modèles pour différentes modalités d'interaction et le prototypage d'une plate-forme "territoire augmenté".

Le projet est structuré en trois étapes :

- La production et l'exploitation visuelle "d'environnements synthétiques" axées sur la numérisation, la structuration et la visualisation de modèles 3D urbains et de leurs contenus associés.
- La production à l'échelle de l'Ile-de-France de bases d'information géo-référencées permettant d'alimenter de

nouveaux services d'information urbains ; le projet vise à être une vitrine en matière de contenus innovants et de valorisation multimédia du patrimoine culturel et touristique.

- La réalisation de briques applicatives et le déploiement de nouveaux produits et services à destination des marchés d'accès aux services urbains, au tourisme, à la valorisation du patrimoine, à la sécurité et à la prévention des risques, aux systèmes d'aide à la décision en temps réel, à la sécurité civile, etc.

Le consortium du projet réunit une vingtaine de partenaires, académiques et industriels. La coordination et l'encadrement sont assurés par Thales Training & Simulation. La structuration technique est organisée en *Workpackages* (WP). Cette thèse s'inscrit dans le WP 3, développement et intégration technologique, et plus précisément dans le WP 3.1.2, Technologies de filtrage et fusion de données, et le WP 3.1.4, Technologies de traitement et de reconstruction.

## 1.2 Architecture haussmannienne

Paris, «Ville Lumière», est aussi une ville riche par rapport à son architecture. Celle-ci doit beaucoup aux travaux de Georges Eugène Haussmann sous le Second Empire. A partir des années 1840, le baron Haussmann fait percer la plupart des axes les plus fréquentés aujourd'hui, tels le Boulevard Saint-Germain, le Boulevard Saint Michel, . . . et le Boulevard Haussmann. L'image de Paris et de ses grands axes est également liée à celle de l'alignement de beaux immeubles d'égale hauteur (dits 'haussmanniens»).

L'architecture haussmannienne ayant été réglementée, les immeubles de cette époque ont des caractéristiques spécifiques. Par exemple, ils ont au moins cinq étages, leurs façades sont en pierre de taille et les murs en brique et serti par des balcons filants aux deuxième et cinquième étages.

Les textes de loi indiquent les valeurs approximatives des dimensions <sup>1</sup> de l'immeuble Haussmannien en fonction de la largeur des voies :

Largeur voie	Hauteur du Gabarit
Inférieure à 7.90m	11.70m
Inférieure à 9.75m	14.62m
Inférieure à 20.00m	17.55m

La façade respecte les dimensions suivantes :

- La hauteur des étages est supérieure à 2.60m (loi 1859).
- La profondeur des saillies (tous les éléments sortant du nu de la façade : balcons, moulures, macarons, etc. ) doit être inférieure à 0.80m (loi 1823).
- Le toit doit s'inscrire sous une diagonale partant de la corniche, à 45° (loi 1784, non appliquée).

1. Ces valeurs ont été extraites du document *Caractérisation des typologies et styles de façades Haussmannienne et Post-Haussmanniennes parisiennes* [77]



- La hauteur du toit est plus ou moins égale à la moitié de la largeur du bâtiment (loi 1784).

### 1.3 Données

Pendant la thèse, plusieurs types de données ont été traités. La description des données sur lesquelles nous avons travaillé est réalisée en fonction du partenaire les ayant fournies : Matis de l'IGN, CAOR de Mines-ParisTech, MENSI – Trimble.

La zone de test choisie pour le projet est une section du 5ème arrondissement de Paris. Elle comprend : le boulevard Saint Michel N° 65-105, la rue Soufflot, la place du Panthéon, la rue Saint Jacques N° 240- 252, la rue Gay Lussac N° 1-32, (Figure 1.1). Cette zone a été choisie parce que c'est une zone caractéristique de l'architecture haussmannienne.

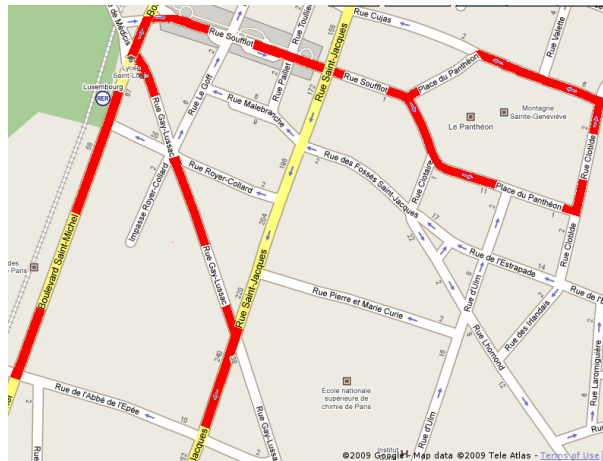


FIGURE 1.1 – Screenshot Google Maps. En rouge la zone de test

#### 1.3.1 Matis de l'IGN

Le laboratoire Matis de l'Institut Géographique National (IGN) a mis à disposition du projet l'ensemble des données suivantes :

##### Images Terrestres et nuages de points

Les images terrestres (2D) et les nuages de points (3D) ont été acquis avec le système mobile STEREOPOLIS. Les données 2D et 3D sont recalées et géo-référencées. Les données 3D fournissent également l'information sur la réflectivité du laser (Figure 1.2). Deux capteurs 3D ont été utilisés avec des angles de numérisation différents : 45° et 90°. En ce qui concerne les images, 12 caméras ont été utilisées avec différents points de vue produisant le

même nombre d'images pour chaque point de prise (Figure 1.3). Chaque image a une taille de  $1920 \times 1080$ .

### Images Aériennes, MNS et Bati3D

Les images aériennes couvrent toute la commune de Paris à une résolution d'environ 10 cm (Figure 1.4(a)). Chaque image a une taille de  $7680 \times 13824$ . L'IGN a également fourni le Modèle Numérique de Surface (MNS) et le Modèle du Bâtiment 3D (Bati3D) produits par des images de résolution de 20cm (Figure 1.4(b) et Figure 1.4(c)).

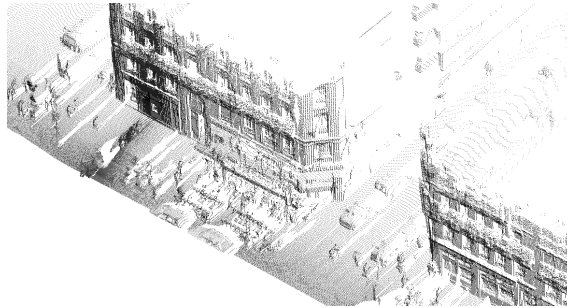


FIGURE 1.2 – Données ©IGN – Rue Soufflot

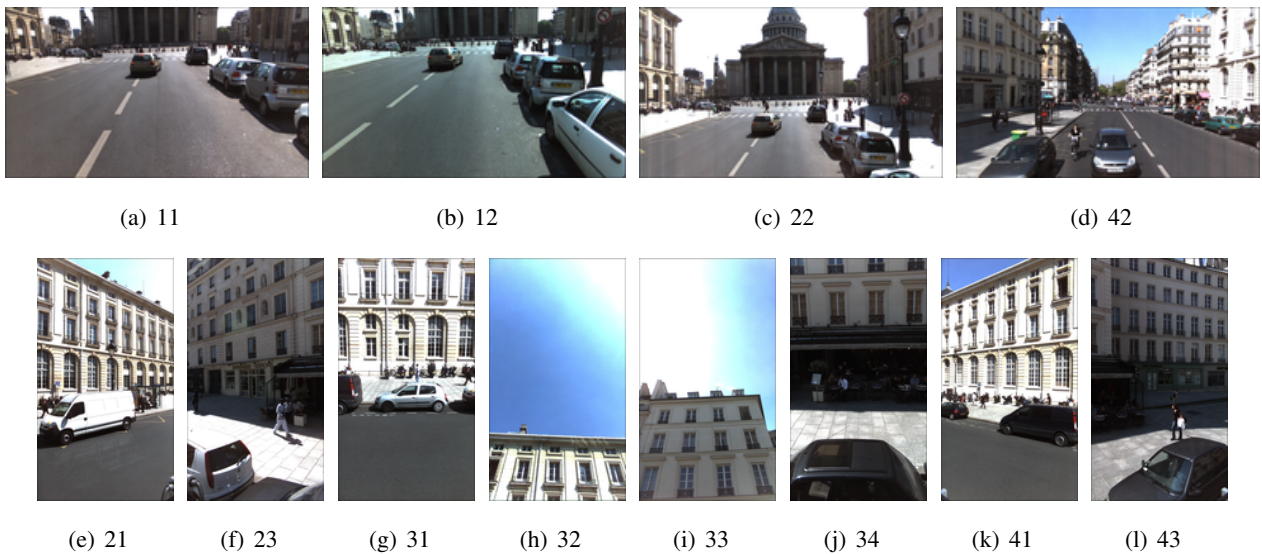


FIGURE 1.3 – Séquences d'images Données ©IGN – Rue Soufflot

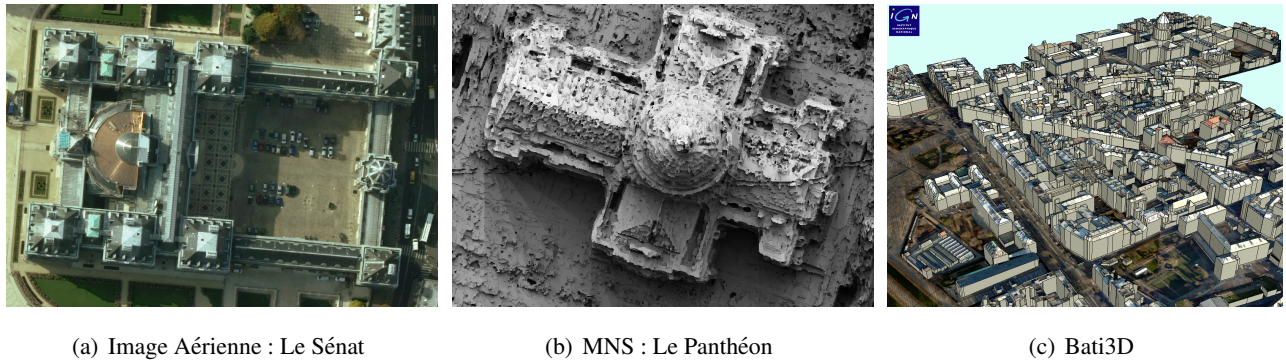


FIGURE 1.4 – Données ©IGN

### 1.3.2 CAOR Mines-ParisTech

Le centre de robotique (CAOR) de l'école Mines-ParisTech a mis à disposition du projet des nuages de points issus de leur système Lara 3D. Ces nuages ont également la couleur de chaque point 3D (Figure 1.5).

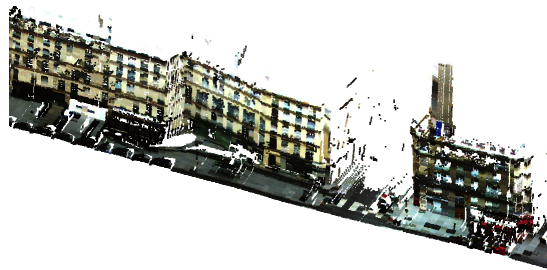


FIGURE 1.5 – Données ©CAOR – Rue Soufflot

### 1.3.3 MENSI - Trimble

MENSI de Trimble a fourni des données de type image et des nuages de points pris au sol avec des systèmes fixes (Figure 1.6). Elles sont recalées et géo-référencées. Les données ont été acquises avec deux systèmes VX et GX. Les données 3D produites par le GX sont hautement denses mais les images 2D sont de basse résolution. Par contre, les données produites par le VX sont des nuages de points moins denses et les images de texture de haute résolution, ainsi que des images prises par un appareil photo classique et calibrées à leurs systèmes.

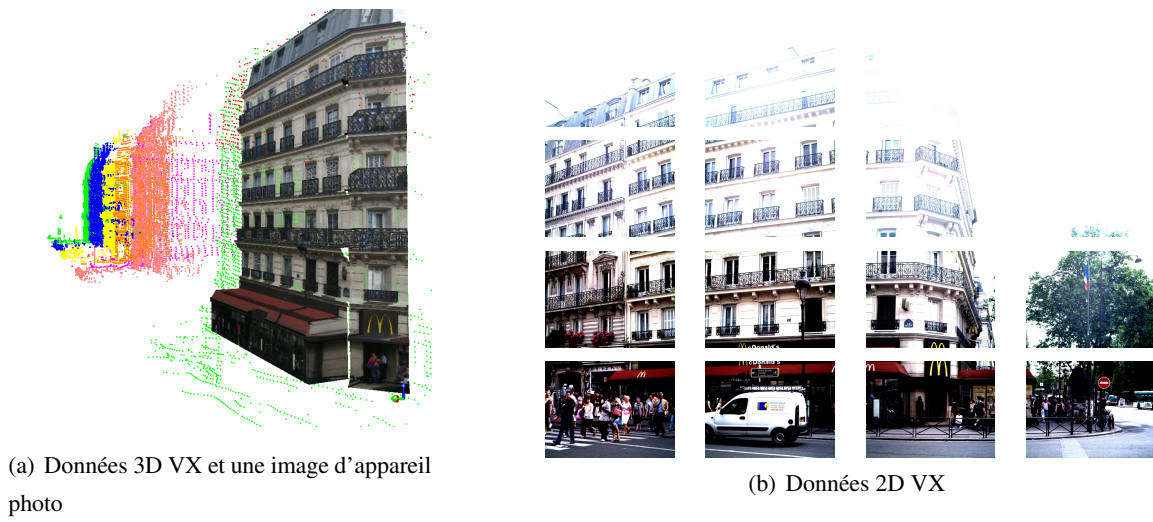


FIGURE 1.6 – Données ©MENSI – Rue Soufflot Bâtiment N 23

---

# **Etat de l'Art : Modélisation d'Environnements Urbains**

---

## **Sommaire**

---

<b>2.1</b>	<b>Introduction</b>	<b>17</b>
<b>2.2</b>	<b>Techniques d'acquisition</b>	<b>18</b>
2.2.1	Images/Vidéos Terrestres	19
2.2.2	Images Aériennes	20
2.2.3	Scanner au sol	21
2.2.4	Scanners aéroportés	22
2.2.5	Sources Dérivées	22
<b>2.3</b>	<b>Techniques de recalage et de filtrage</b>	<b>23</b>
2.3.1	Techniques de recalage	23
2.3.2	Techniques de filtrage	24
<b>2.4</b>	<b>Techniques d'infographie</b>	<b>25</b>
2.4.1	Modélisation manuelle	25
2.4.2	Modélisation procédurale	26
<b>2.5</b>	<b>Techniques de reconstruction</b>	<b>27</b>
2.5.1	Reconstruction par photogrammétrie	28
2.5.2	Reconstruction à partir des nuages de points	29
2.5.3	Modélisation coopérative	29
<b>2.6</b>	<b>Discussion</b>	<b>31</b>

---

## **2.1 Introduction**

Dans ce chapitre, une étude de la bibliographie concernant la modélisation d'environnements urbains est présentée. Plusieurs critères peuvent être utilisés pour classer les différentes approches de modélisation : les sources d'acquisition de données, le niveau d'interaction avec l'utilisateur, les étapes de modélisation, la fidélité géométrique et les applications prévues. D'abord, nous avons fait un premier classement général par rapport au type de modèle généré : géo-spécifique et géo-typique. Ensuite, chaque sous-groupe est divisé selon les différentes techniques qui

contribuent à la modélisation.

**Définition 1.** *Un modèle **géo-spécifique** est une représentation synthétique de l'environnement avec des contraintes importantes de fidélité à la réalité en termes de localisation, de taille, de couleur, de matériaux, etc... Le modèle s'appuie sur des données prises sur le terrain pour garantir un niveau de fidélité élevé.*

**Définition 2.** *Un modèle **géo-typique** est une représentation synthétique de l'environnement qui doit sembler réaliste mais sans contraintes de fidélité à la réalité elle-même. Par exemple, une rue qui ressemble au style parisien mais sans être une rue, existant réellement à Paris.*

Les modèles géo-typiques sont utilisés dans l'industrie du cinéma ou les jeux vidéo alors que les modèles géo-spécifiques sont nécessaires pour des simulations militaires, du tourisme virtuel ou encore l'aménagement du territoire. Les techniques impliquées dans la création d'un modèle géo-spécifique sont :

- Les techniques d'acquisition comprennent les domaines suivants : les capteurs utilisés, les types de données générées et les localisations d'acquisitions.
- Les techniques de recalage correspondent à la mise en cohérence de diverses sources et de divers types de données.
- Les techniques de filtrage consistent à nettoyer, lisser les données et/ou éliminer le bruit dû à la numérisation.
- Les techniques de reconstruction comprennent l'extraction de la sémantique ou des structures géométriques des données réelles issues des étapes précédentes.

Par ailleurs, les modèles géo-typiques sont produits par des techniques d'infographie. Les approches par infographie/graphisme sont obtenues par des méthodes purement manuelles en utilisant des logiciels spécifiques à la création de modèles synthétiques. Concernant les techniques de modélisation procédurale, ce sont en général des techniques d'infographie pour créer des modèles 3D ainsi que des textures à partir de séries de règles [52]. Ces approches, comme les approches d'infographie pure, peuvent produire de manière indépendante des modèles urbains où la sortie est complètement virtuelle. Néanmoins, de plus en plus de chercheurs combinent l'information acquise avec la modélisation procédurale, donnant pour résultat des modèles géo-typiques plus fidèles au monde réel. Nous les appellerons ces techniques modélisations coopératives. Le classement global présenté est illustré sur le diagramme de la Figure 2.1.

## 2.2 Techniques d'acquisition

La classification des travaux réalisés dans le domaine de l'acquisition des données est faite selon le système utilisé : capteurs passifs, capteurs actifs et sources dérivées. Les capteurs passifs sont principalement les capteurs CCD qui acquièrent l'information de luminance et/ou la couleur d'une scène urbaine sous forme d'images ou de vidéos. Pour extraire l'information 3D à partir des images/vidéos, l'information de calibration de la caméra est nécessaire.

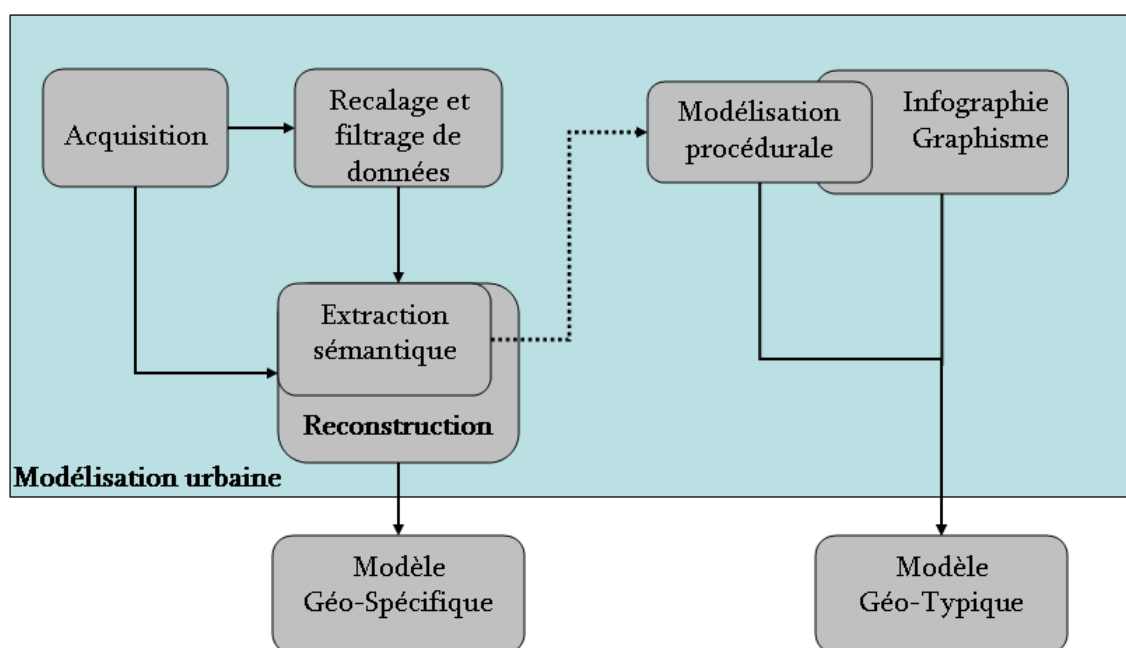


FIGURE 2.1 – Les étapes de la modélisation d'environnements urbains.

Les conditions d'acquisition jouent un rôle important pour ce type de capteurs car les conditions d'éclairage ne sont pas contrôlées. Par ailleurs, les capteurs actifs mesurent directement la profondeur des objets et fournissent sous forme de nuages de points la géométrie de la scène pour la modélisation d'environnements urbains. Les techniques les plus utilisées pour l'acquisition des données sont le temps de vol et la triangulation optique. Les capteurs passifs ou actifs peuvent être divisés en deux sous catégories selon le point d'acquisition : terrestres et aéroportés. Vers la fin des années 80, les chercheurs ont commencé à employer de nouveaux systèmes de navigation de géo-positionnement et de guidage (GPS/INS *Global Positioning System /Inertial navigation system* ) pour le positionnement précis de capteurs. La combinaison des deux systèmes a augmenté la précision et a diminué le coût d'obtention d'une grande quantité de données géo-référencées. Finalement, il existe une grande variété de sources dérivées : le cadastre, le Modèle Numérique d'Élévation (MNE), les Systèmes d'Information Géographique (SIG). Les approches mentionnées précédemment sont détaillées dans les sections suivantes.

### 2.2.1 Images/Vidéos Terrestres

Les images terrestres sont acquises au niveau du sol à l'aide de systèmes fixes. Bien que les données fournies soient de qualité élevée (végétation, détail de façade de bâtiment, etc.), il manque des informations sur la vue supérieure des bâtiments (toits) et les occlusions rendent difficile l'extraction de l'information 3D. Il est donc difficile à partir de ces données de construire de grands secteurs urbains. Pour cette raison, la modélisation à partir

d'images terrestres est plus utile dans les applications géo-typiques fournissant des modèles 3D de qualité [8, 168]. En effet, l'estimation du modèle 3D texturé avec l'utilisation d'images de textures réelles permet d'augmenter son réalisme visuel.

La photogrammétrie présente aussi des problèmes affectant la qualité dans l'étape de reconstruction, par exemple [170] :

- Les problèmes de distorsion dus à la perspective pendant l'acquisition.
- Les variations d'éclairage qui produisent différentes luminances et des ombres dans l'image.
- Les occlusions provoquées par les objets à modéliser (autres bâtiments) ou les objets à ne pas modéliser tels que des arbres, des poteaux de service, des piétons et des voitures.

Dans les séquences vidéo, la résolution des images/frames n'est pas aussi élevée que dans le cas des images fixes. Pollefeys a proposé des techniques pour récupérer automatiquement la géométrie 3D en utilisant un simple caméscope portable [125, 126]. Cette approche présente une application de modélisation des scènes archéologiques qui ne requiert pas de grands déplacements. Pour acquérir de larges zones urbaines, les capteurs sont normalement situés sur des plateformes mobiles. Dans le projet UrbanScape [3, 127], un système d'acquisition sur un véhicule embarqué avec huit caméras vidéo et un système de navigation INS/GPS a été développé. Brun *et al.* utilisent un objectif *fisheye* afin d'obtenir des images grand-angle [25].

### 2.2.2 Images Aériennes

La modélisation à partir d'images aériennes présente plusieurs avantages par rapport à la modélisation à partir d'images terrestres. Les images aériennes fournissent les formes de bâtiments et les dimensions précises de toits. Par contre, elles n'ont pas une information suffisante au niveau des façades. Les images peuvent être rectifiées dans la projection orthographique, facilitant la fusion des multiples images pour couvrir de grands domaines géographiques. Ces types de données facilitent également la fusion avec les SIGs. La modélisation avec ces images présente les mêmes contraintes qu'avec des images terrestres par rapport aux conditions d'éclairage. Plusieurs systèmes d'acquisition d'images aériennes et leur utilisation pour la reconstruction de bâtiments 3D sont décrits dans [57]. Maillet et Flamanc illustrent une comparaison de l'acquisition des images aériennes et des images satellites [98]. Chaque type d'image offre différentes caractéristiques parmi lesquelles la qualité, la précision, la couverture, le coût et les installations opérationnelles. Les images aériennes sont celles qui présentent le meilleur rapport qualité/couverture.



### 2.2.3 Scanner au sol

Les systèmes de capteurs actifs utilisent la technologie Lidar pour l'acquisition de données. Lidar est l'acronyme du terme anglo-saxon *light detection and ranging*. C'est un système de mesure utilisant des ondes électromagnétiques. Le laser émet une onde lumineuse. L'onde interagit avec les différents composants qu'elle rencontre. Une partie de cette onde est rétrodiffusée et collectée par le capteur. À partir de cette composante rétrodiffusée, on peut alors déduire des informations quant au diffuseur et sa distance par rapport au système de mesure.

Früh et Zakhorsky ont présenté un système qui utilise un véhicule équipé d'un appareil photo et de deux scanners laser 2D (horizontal et vertical) [47]. L'appareil photo capture des images pour les textures, le scanner laser horizontal suit le mouvement du véhicule, et le scanner vertical capture des données de la façade des bâtiments. Zhao et Shibasaki ont employé un système de véhicule avec des capteurs embarqués : trois scanners laser 1D et six appareils photo alignés [181]. Le système de navigation fusionne les données du GPS, de la centrale inertielle, et de l'odomètre. Ce système permet de géo-référencer et de synchroniser les nuages de points et les images. De même, le laboratoire de recherche CAOR de Mines-ParisTech a développé un véhicule pour l'acquisition de données 3D géo-référencées [2, 24]. Ce véhicule prototype, Lara3D, comprend principalement des capteurs de positionnement (GPS différentiel, centrale inertielle, odomètre) pour assurer la localisation spatiale. Un scanner laser 2D embarqué permet la numérisation d'environnements urbains et routiers, au cours du déplacement du véhicule. Un système mobile sur véhicule avec des caractéristiques similaires a été développé par le laboratoire MATIS de l'IGN (Institut Géographique National de France). Ce système, nommé STEREOPOPOLIS, possède aussi deux caméras de haute résolution, un laser 2D et un système de navigation pour le géo-référencement.

Il existe d'autres types de capteurs actifs 3D au sol. Ces types de capteurs sont fixes et de haute gamme, et du fait d'un prix élevé, ils sont seulement produits pour des applications industrielles ou de construction. Par exemple, MENSIS-Trimble®<sup>1</sup> possède deux types de dispositifs pour l'acquisition de nuages de points :

- La station spatiale Trimble® VX<sup>TM</sup> est un système de positionnement évolué qui fait appel aux technologies de pointe de numérisation et d'optique pour réaliser des mesures de faible résolution 3D et générer des données métriques et géo-référencées en 2D et 3D.
- Les scanners Trimble® 3D GX emploient la technologie de mesure du temps de vol dont le principe est d'envoyer une impulsion laser et d'observer le temps qu'elle prend pour se refléter sur un objet et pour retourner à l'instrument. La gamme de distance est combinée avec des mesures d'encodeur d'angle pour fournir la localisation tridimensionnelle d'un point. Ces types de système fournissent des données plus denses que les systèmes précédemment présentés.

---

1. <http://www.trimble.com/spatialimaging.shtml>

### **2.2.4 Scanners aéroportés**

Dans le cas des systèmes aéroportés avec des capteurs actifs, le laser est embarqué dans un porteur aérien (avion, hélicoptère, dirigeable, deltaplane, etc). Il émet des impulsions vers la terre et le système acquiert un nuage de points géo-référencés du terrain [60]. Un état de l'art étendu des systèmes aéroportés Lidar est présenté en [100]. Les auteurs décrivent les principes physiques de la technique et les systèmes existants les plus importants.

### **2.2.5 Sources Dérivées**

#### **Modèle Numérique d'Élévation (MNE)**

Le MNE est normalement produit par stéréoscopie à partir de couples d'images aériennes donnant pour résultat un modèle de la topographie d'une zone terrestre. Ce modèle d'élévation est représenté comme une image où la valeur de chaque pixel correspond à l'altitude d'un point sur le terrain.

#### **Cadastre**

Le cadastre est le registre public et les documents administratifs (état de section, plan parcellaire, matrice cadastrale) permettant d'identifier les propriétés présentes dans une commune. Le cadastre recense toutes les propriétés foncières.

#### **Bati3D**

Le Bati3D est un modèle polyédrique de la structure des bâtiments. Il est issu de l'extraction d'information de la hauteur dans le MNE et la délimitation de l'information cadastrale. Les approches utilisées pour sa création seront décrites dans la section 2.5.1.

#### **Système d'information géographique (SIG)**

Un système d'information géographique peut être défini comme un ensemble de matériels et de logiciels capables d'intégrer, de stocker, d'éditer, d'analyser, de partager, et de montrer l'information géographiquement référencée [29]. Les données graphiques localisées aident à afficher ou à imprimer des plans et des cartes. Ses usages couvrent les activités géomatiques de traitement et de diffusion de l'information géographique. Cela constitue le moyen d'associer différents types d'acteurs, publics et privés, et d'établir des partenariats locaux tout en favorisant une gestion plus décentralisée. Le système est généré à partir de plusieurs sources d'information intégrées typiquement par un utilisateur. Le rôle du système d'information est de proposer une représentation plus ou moins réaliste de l'environnement spatial en se basant sur des primitives graphiques telles que des points, des vecteurs (arcs), des

polygones ou des réseaux (raster) [38]. À ces primitives sont associées des informations qualitatives telles que la nature (route, voie ferrée, forêt, etc.) ou des informations contextuelles.

## 2.3 Techniques de recalage et de filtrage

Cette section illustre des techniques de recalage et de filtrage de plusieurs données hétérogènes mentionnées dans la section précédente.

### 2.3.1 Techniques de recalage

L'idée principale du recalage est la mise en correspondance dans le même repère des différents types de données. Les méthodes de recalage dans la modélisation urbaine généralement procèdent en deux étapes : la mise en correspondance des données et l'estimation de la transformation permettant de projeter toutes les données dans un même repère. Il faut considérer que si les données sont acquises dans le même instant de temps avec des capteurs complètement synchronisés, le problème du recalage est réduit à un problème d'instrumentation du matériel. Les techniques de recalage peuvent être classées par rapport à plusieurs critères : les applications, le type de données, le type de transformation, le type d'optimisation, etc. Nous décrirons ces techniques en utilisant un classement similaire à celui présenté par Stamos dans [153] par rapport aux types de données.

#### 3D-3D

La méthode la plus répandue est l'algorithme ICP (Iterative Closest Point), développée par Besl et McKay [11] et Zhang [180]. Cet algorithme consiste à calculer, de façon itérative, la matrice de transformation recalant au mieux deux (ou plusieurs) ensembles de données 3D. Des nombreuses variantes et améliorations de cet algorithme existent [1].

#### 2D - 3D

La plupart des travaux orientés vers le recalage de nuages de points et d'images prises au sol sont basés sur l'extraction de primitives dans les données (droites [154], rectangles [155], plans [36], régions [32], etc.) et leur mise en correspondance *a posteriori*. D'autres approches réduisent le problème du recalage multimodal 2D-3D à un problème 2D-2D [33, 145] en utilisant l'image de reflectance associée au capteur 3D. La mise en correspondance est alors réalisée entre l'image de reflectance et l'image couleur. Huber a proposé une méthode semi-automatique pour le recalage de données aéroportées [75]. A partir d'une estimation des plans de toits avec les données laser et des contours dans l'image, il met en correspondance les données. D'autres types d'approches recalent des modèles 3D de bâtiments avec des images aériennes [46] ou terrestres [85]. Ils extraient les lignes de fuite dans les images

et comparent avec les contours du modèle 3D. Actuellement, la plupart des fournisseurs de données 3D (fixes et aéroportées) proposent en complément des données laser un ensemble de photographies acquises à partir d'une caméra calibrée avec une bonne précision.

## 2D - 2D

Les différentes familles de recalage d'images que l'on trouve dans la littérature peuvent être divisées en trois grands groupes : recalage iconique, recalage géométrique, recalage hybride [23, 99]. Le recalage géométrique s'appuie sur l'appariement de primitives géométriques extraites des images à recaler. Le recalage iconique ne requiert aucune segmentation et s'appuie directement sur les intensités des images à recaler. Le recalage hybride est une méthode hybride entre les deux précédentes. Dans les images de milieu urbain, les approches les plus répandues sont issues des méthodes hybrides basées sur la détection de points d'intérêt. Dans la littérature on trouve de nombreuses méthodes de détection des points d'intérêt. Un état de l'art complet sur la détection des points d'intérêt est présenté dans [110]. Lindeberg a développé un détecteur de «blob» invariant aux facteurs d'échelle, où un blob est défini comme un maxima du Laplacien normalisé dans l'espace des échelles [90]. Par la suite Lowe a réalisé une approximation du Laplacien à partir de filtres successifs basés sur des différences de gaussienne (DoG), il détecte les maxima locaux dans l'espace des échelles [94,95]. Mikolajczyk et Schmid ont utilisé une version multi échelle du détecteur de points d'intérêt de Harris pour localiser les points singuliers de l'image dans l'espace, puis la méthode de Lindeberg pour la détection des échelles et l'adaptation aux transformations affines [109].

Par ailleurs, des méthodes de recalage d'images aériennes et terrestres en utilisant un modèle polyédrique sont présentées dans [74, 86]. Les images aériennes sont d'abord utilisées pour recréer le modèle et ensuite les images terrestres sont assemblées au modèle.

### 2.3.2 Techniques de filtrage

Les données brutes, issues de l'étape d'acquisition, contiennent toutes sortes d'informations liées à l'environnement urbain. Cependant, pour faciliter la modélisation, il est intéressant de seulement avoir les données correspondant aux structures d'intérêt. Diverses approches ont développé des méthodes de segmentation pour ensuite sélectionner les «régions» à modéliser. Les approches de filtrage sont décrites selon les types de données sur lesquelles elles sont appliquées.

#### Images

Les méthodes de filtrage sur des images sont normalement appliquées à des images prises au sol, où le problème des occlusions est important (des arbres, des panneaux, des personnes). Le filtrage est principalement composé de deux étapes : tout d'abord, la sélection de pixels d'éléments à filtrer (segmentation) et ensuite la reconstruc-

tion de l'information sélectionnée (*Inpainting*). Wang *et al.* ont réduit les zones occultées en utilisant plusieurs images [170]. Ils pondèrent les pixels de toutes les images avec un poids qui dépend de l'orientation de la caméra et de la corrélation entre images. Koran et Rasmussen ont présenté une méthode de classification de pixels d'obstacles et de non-obstacles [80]. La méthode recalcule en même temps les séquences d'images en utilisant une extraction de caractéristiques et RANSAC (*RANdom SAMple Consensus* [42]). Les trous, produits par l'élimination de pixels occultés, sont remplacés avec une estimation spatio-temporelle de la luminance des pixels. Une détection de voitures statiques dans des séquences vidéo a été développée en [28]. L'approche est basée sur l'apprentissage a priori avec des caractéristiques de voitures et l'hypothèse de la localisation de voitures au niveau du sol.

### Nuages de points

Madhavan et Hong détectent la séparation de bâtiments et de routes sur des données Lidar Terrestres [97]. Ils séparent également en plusieurs pâtés de maisons si les données comportent plusieurs îlots. Goulette et al. présentent une segmentation de nuages de points basée sur l'analyse de profils de points et leur accumulation [56]. Ils détectent la route, les façades et les arbres. Néanmoins, ces dernières approches ont de fortes contraintes par rapport à leurs systèmes d'acquisition. Des approches de détection de véhicules sur des données aériennes sont décrites en [30, 175]. Ces approches présentent des problèmes de détection sur des routes en pente. On pourrait aussi mentionner les approches qui modélisent des plans de façades à partir des nuages de points [9, 18, 37, 156, 177] pour éliminer les points aberrants, mais elles seront traitées comme des méthodes de reconstruction (voir section 2.5.2).

## 2.4 Techniques d'infographie

### 2.4.1 Modélisation manuelle

En effet, la modélisation manuelle par infographie se base sur un modèle de synthèse entièrement généré par ordinateur dans laquelle le graphisme est utilisé pour créer un monde visuellement réaliste. Cette approche est un sous-domaine de la Conception Assistée par Ordinateur (CAO). L'utilisation d'un logiciel suffisamment puissant pour la création des modèles synthétiques est nécessaire, parmi lesquels nous pouvons citer 3DSMax<sup>2</sup>, Maya<sup>3</sup>, AutoCAD<sup>4</sup>, Rhinoceros<sup>5</sup> et Blender<sup>6</sup>. Les applications sont nombreuses, mais les plus remarquables sont des jeux vidéo, des logiciels d'architecture urbaniste, des films animés, etc. [123].

Ce type de modélisation présente divers verrous décrits dans [53] :

---

2. [www.autodesk.es/3dsmax](http://www.autodesk.es/3dsmax)

3. [www.autodesk.com/maya](http://www.autodesk.com/maya)

4. [www.autodesk.com/autocad](http://www.autodesk.com/autocad)

5. [www.rhino3d.com/](http://www.rhino3d.com/)

6. [www.blender.org/](http://www.blender.org/)

- Le coût de production : lié au nombre élevé d'années-hommes dans l'industrie du jeu et du cinéma. Par exemple 15 années-hommes pour la modélisation de *Metropolis* dans le film *Superman Returns* [113] et 10 années-hommes pour la modélisation de New York années 30 dans le film *King Kong* <sup>7</sup>.
- La durabilité et la réutilisation : le cycle de vie court et une réutilisation postérieure nulle des modèles créés.
- L'expertise : l'utilisation des logiciels requiert des mois de formation.

Par contre, la modélisation procédurale propose une alternative intéressante pour réduire le coût de production, augmenter la généralité des modèles générés et simplifier la modélisation comme un jeu de Lego.

## 2.4.2 Modélisation procédurale

Les techniques de modélisation procédurale facilitent la production automatique des scènes/modèles complexes. Elle garantit aussi une cohérence globale dans le modèle et un très haut niveau de réalisme. Une des caractéristiques les plus importantes des techniques de modélisation procédurale est le niveau d'abstraction. Dans une approche procédurale, plutôt que de préciser explicitement et de stocker tous les détails complexes d'une scène ou d'un modèle, le modèle est synthétisé dans une fonction ou un algorithme (c'est-à-dire une procédure) [40]. L'approche économise la mémoire de stockage, car les détails ne sont plus explicitement spécifiés, mais sont implicites à la procédure. Cela permet de créer des modèles multi-inhérents, à différents niveaux de détails que l'on peut évaluer à une résolution déterminée. Dans la modélisation d'environnements urbains, la procédure est basée sur des règles (constituant une grammaire). Différents niveaux de modélisation procédurale ont été présentés dans [122] en utilisant une approche multi-échelle. Cette modélisation produit principalement des modèles géo-typiques.

Les premières approches basées sur une grammaire [148] [128], y compris graftals et L-systèmes, permettent de simuler des modèles complexes d'arbres, de plantes, et d'autres objets naturels à partir de la définition de quelques paramètres. Ces modèles utilisent des langages formels pour spécifier des règles complexes dans la croissance naturelle des objets.

La génération de milieux urbains étendus (routes, îlots et bâtiments) a été présentée par Parish et Müller [120], Lechner *et al.* [84] et Marvie *et al.* [102]. Ils utilisent un L-système étendu pour faire "pousser" les routes comme des plantes. Les zones entre les routes sont ensuite subdivisées pour définir les îlots de bâtiments. Les bâtiments sont générés avec un autre L-système et composés sous des modèles polyédriques. Mais les L-systèmes ne sont pas nécessairement appropriés pour la modélisation des bâtiments. Les bâtiments diffèrent dans la structure des plantes et des rues, dans la mesure où ils ne poussent pas de manière libre, leur modélisation repose sur une partition des espaces plutôt que sur un processus de croissance. Ainsi, Wonka *et al.* ont développé une méthode automatique pour la modélisation qui permet de reconstruire les différents types de bâtiments en utilisant un ensemble de règles. Ils ont montré comment générer des détails géométriques sur les façades de bâtiments à partir de grammaires de formes (*shape grammars*) [157]. Ces grammaires utilisent des formes géométriques (*shapes*) comme un alphabet.

---

7. [www.wetafx.co.nz/](http://www.wetafx.co.nz/)

Ils proposent deux types de règles principalement : les grammaires de sous-division (*split*), qui divisent le bâtiment en parties, et les grammaires de contrôle, qui guident la propagation et la distribution des formes. Leurs résultats présentent une diversité culturelle et répondent à l'influence de la population et du matériel. Les règles de Wonka *et al.* ont été étendues par Müller *et al.* en [113]. Ils ont nommé la grammaire CGA *shape* (acronyme du terme anglo-saxon *Computer Graphics Architecture*). Ils permettent de spécifier les paramètres et emploient une grammaire et des attributs prédéterminés afin de générer automatiquement des bâtiments et des architectures. Une des contributions de ces travaux concerne la modélisation des bâtiments à partir d'un assemblage de volumes simples. Ces deux approches font partie du logiciel *City-Engine*<sup>8</sup> pour la modélisation d'environnements urbains. Aliaga *et al.* décrivent un système similaire qui facilite l'utilisation de grammaires pour la description de façades [6].

Un autre type de modélisation procédurale a été introduit par Havemann dans sa thèse doctorale [62]. Sa méthode est appelée GML (acronyme du terme anglo-saxon *Generative Modeling Language*). Le GML est comme un langage de programmation, qui permet de créer les détails basés sur des maillages en utilisant des grammaires et de générer des surfaces complexes à partir de plusieurs surfaces simples. Néanmoins, la grammaire n'est pas aussi bien adaptée pour des ornements et des surfaces courbes.

Certaines applications utilisent la modélisation procédurale dans le domaine du cinéma [171], de l'aménagement urbain [52, 111], de l'infrastructure souterraine [106], de la conservation du patrimoine : les maisons d'Alvaro Siza à Malagueira [39], des constructions du Moyen Âge [16], des châteaux [53], des temples chinois [91], et des bâtiments Maya [112].

## 2.5 Techniques de reconstruction

La modélisation d'environnements urbains réels à grande échelle est réalisée à partir de données issues des approches déjà présentées où le but est de produire des modèles géo-spécifiques. Chacune de ces données et techniques d'acquisition ont des avantages et des inconvénients au moment de la modélisation. Par exemple, les images apportent des informations de couleur et texture de haute qualité et ne fournissent pas de données 3D, par contre les capteurs actifs donnent l'information 3D avec une précision élevée sans information couleur. Par ailleurs, en fusionnant les différents types de données entre eux avec la modélisation procédurale, on pourrait produire des modèles urbains plus précis et d'une manière plus automatique. Dans les techniques de reconstruction, on parlera également des méthodes qui détectent la sémantique (étages, travées, fenêtres) et qui alimentent une grammaire déterminée.

---

8. [www.procedural.com/cityengine/](http://www.procedural.com/cityengine/)

### 2.5.1 Reconstruction par photogrammétrie

La modélisation à partir de vidéos/images est un problème classique dans la vision par ordinateur [45]. Ces techniques emploient seulement des images pour obtenir des modèles 3D ; pour cette raison, c'est une manière rentable de produire des modèles urbains à grande échelle [73]. La modélisation est seulement possible si plus de deux images ont une information commune, cette information est utilisée pour trouver des correspondances et finalement, grâce aux techniques de stéréogrammétrie, on obtient des données 3D. Vu *et al.* utilisent des images de haute résolution afin d'extraire l'information 3D des zones étendues [168]. Ils utilisent la programmation sur le processeur graphique d'une machine pour garder un temps raisonnable d'exécution.

L'un des premiers travaux par photogrammétrie terrestre est présenté par Debevec *et al.* [31]. Le système obtient des modèles géométriques à partir d'images de façades acquises depuis un système terrestre. L'information de façades est exploitée en utilisant l'extraction des primitives de chaque image. Avec un système interactif, l'utilisateur choisit les primitives pour établir les correspondances. La précision du modèle obtenu est vérifiée en projetant à nouveau le modèle sur les images originales. De la même manière, les séquences vidéo d'une scène statique, acquise par une caméra en mouvement, sont utilisées pour faire la reconstruction de modèles géométriques 3D [125, 126].

Lin *et al.* [89] emploient des images aériennes monoculaires d'un point de vue général pour détecter des bâtiments urbains et pour construire des descriptions de forme 3D. Ils extraient la forme des bâtiments à partir de deux hypothèses : les toits sont plats et rectangulaires et les bâtiments ont seulement des formes en  $L$ ,  $T$ , et  $I$ . Baillard [7] a décrit deux approches de reconstruction automatique des bâtiments en se basant sur deux images aériennes. La première approche est un algorithme pour trouver automatiquement les correspondances entre des segments de ligne de multiples images. L'algorithme utilise des contraintes géométriques et radiométriques basées sur l'information de vues multiples. Le deuxième développement est une méthode pour calculer automatiquement une reconstruction par morceaux planaires basée sur les correspondances des lignes. Le résultat est un modèle brut de bas niveau de détail où la forme du toit est un carré (le premier bati3D). La mise en correspondance d'images donne lieu au modèle numérique d'élévation. Ce travail est le début d'une série de thèses de doctorat [7, 50, 119, 158] dans le Laboratoire de recherche MATIS de l'IGN. Dans sa thèse, Taillandier [158] a utilisé six images aériennes pour trouver les correspondances et les modèles obtenus ont plus de détails au niveau de toits que dans l'approche de Baillard. Il a obtenu un modèle de bati3D plus élaboré avec de toits en pente. Récemment, Brédif *et al.* introduisent les superstructures au toit [19]. Ces superstructures sont des détails géométriques comme des cheminées, des chiens-assis, des verrières, des terrasses de toits.

Dans les dernières années, une autre approche présente la modélisation 4D d'une base de données photos qui montre l'évolution dans le temps d'une ville. Cette approche a été développée dans le cadre du projet 4D Cities [141, 142]. La base de données d'images correspond à une période de 100-150 ans pour recréer des modèles 4D (modèles 3D de bâtiments + temps) utilisant également des techniques de mouvement.



### 2.5.2 Reconstruction à partir des nuages de points

La plupart des recherches sur l'exploitation de l'information 3D de lidar terrestres concerne l'approximation des façades par des plans. Dans [18,37], des méthodes de croissance de régions sont utilisées pour extraire les surfaces planaires de façades et dans [9,156] l'approximation de la façade est réalisée en utilisant l'algorithme de RANSAC. Becker et Haala ont détecté la position de fenêtres pour déduction, puisque le laser traverse la vitre, il n'y a pas de points aux emplacements des fenêtres [9]. Yu *et al.* ont présenté un algorithme de segmentation qui extrait les plans d'un ensemble de nuages de points recalés [177]. La méthode utilise une variante de l'algorithme d'agroupement *k-means*. Le résultat est un nombre de régions qui décrivent les entités urbaines (par exemple, les façades, les fenêtres, les plafonds, les détails architecturaux, etc).

Par ailleurs, les données lidar aériennes sont utilisées pour la création des modèles polyédriques de bâtiments [79, 144,163]. Verma *et al.* segmentent le nuage 3D en points de toiture (bâtiment) et en points de terrain. Ils introduisent le concept d'une topologie de toit basée sur des graphes pour représenter les relations entre les différents plans d'une structure complexe de toiture. Cette représentation permet la composition de la toiture avec la combinaison de simples plans.

Toutes ces techniques de reconstruction utilisées avec les données lidar sont aussi appliquées aux données 3D issues des méthodes de stéréogrammétrie.

### 2.5.3 Modélisation coopérative

Früh et Zakhor ont présenté la modélisation des bâtiments en utilisant l'information d'un modèle aéroporté et des images terrestres [48, 49]. Ils génèrent des modèles triangulés avec des texels de textures associés. Hu *et al.* ont employé des données de Lidar, une image aérienne pour reconstruire des modèles polyédriques de bâtiments [73]. Ces modèles sont texturés avec des images terrestres. Vosselman and Dijkman ont utilisé des modèles MNE ou des images aériennes et des SIG 2D pour la reconstruction des modèles géo-référencés [167]. Bretar [21] a étudié l'utilisation conjointe de données laser avec des images dans une approche par segmentation hiérarchique, dans le cadre de reconstructions de bâtiments à partir d'acquisitions aéroportées.

Les approches les plus représentatives d'une modélisation coopérative ont pour but la reconstruction de façades des bâtiments. Wang *et al.* ont présenté un algorithme qui récupère les structures détaillées des façades de bâtiments à partir d'images de multiples vues [170]. Elles doivent être ortho-rectifiées. Cet algorithme détecte une classe générique d'objets (microstructures) montrant une taille, un modèle, et une orientation régulière (*Oriented Region Growing* (ORG) [169]). Après cela, ils considèrent que les microstructures de tailles semblables ont un modèle périodique dans des directions horizontales et verticales sur la façade (*Periodic Pattern Fixing* (PPF)). Werner et Zisserman ont décrit une méthode pour reconstruire la structure détaillée d'un bâtiment et de ses façades en adaptant des blocs primitifs [172]. Ces blocs sont automatiquement adaptés afin de reconstruire les fenêtres du toit.

Ils emploient la géométrie projective sur l'image non calibrée. D'abord, un modèle extérieur brut du bâtiment est construit en utilisant la fusion des images avec des techniques de stéréogrammétrie. Ensuite, le modèle brut est créé, et affiné d'extérieur en ajoutant des détails tels que les fenêtres et les portes. La reconstruction emploie la détection des points de fuite et la correspondance des lignes qui sont nécessaires pour la détection des primitives planaires. Schindler et Bauer ont proposé une méthode de reconstruction détaillée des bâtiments basée sur des modèles [143]. Dans cette reconstruction, les points 3D sont obtenus en trouvant les correspondances entre les lignes des images. Ces points sont utilisés pour créer un modèle brut polyédrique avec un algorithme robuste de régression. Les détails des façades sont générés à partir d'images couleur terrestres.

Parish et Müller [120] construisent des modèles de villes en utilisant des images aériennes et l'information géographique de la densité de population. Ils utilisent également un L-Système étendu avec des contraintes. Un utilisateur peut déterminer le modèle des axes routiers : radiaux comme Paris ou rectangulaires comme New York, etc... Müller *et al.* ont récemment proposé une méthode basée sur des images de façades, pour reconstruire de manière procédurale des modèles de bâtiments [114]. Cette solution combine la modélisation procédurale avec l'analyse d'images pour obtenir un partitionnement hiérarchique de façades ortho-rectifiées. Les façades sont d'abord divisées par étages et ensuite par travées. Les travées sont sous-divisées en fenêtres et murs. Cet algorithme est basé sur une détection probabiliste, information mutuelle, de la symétrie dans les images de façades sur les coordonnées X et Y pour obtenir une réduction des images appelées façades irréductibles. L'approche est étendue avec des images de façades de forte perspective par Van Gool *et al.* [55]. La méthode est basée sur une analyse de points de fuite et une similitude entre caractéristiques pour trouver la structure répétitive de la façade. Une méthode interactive a été présentée par Hohmann *et al.* [72] pour la modélisation des bâtiments de la ville de Graz dans le cadre du projet CityFit. Ils utilisent le langage GML pour décrire la grammaire et l'information conjointe de données images et laser pour aider l'utilisateur dans l'extraction de l'information sémantique.

Un sujet important pour la reconstruction automatique de façades est la détection de tous ces éléments constitutifs, principalement les fenêtres. Lee et Nevatia, en utilisant une caméra calibrée, extraient et reconstruisent la structure des fenêtres 3D du bâtiment [87]. Ils trouvent automatiquement les fenêtres (des rectangles dans l'image rectifiée) employant une méthode de projection de profil de l'image de gradient, qui exploite la régularité du placement vertical et horizontal des fenêtres. A partir des contours de fenêtres obtenus, et grâce à la calibration de caméra, une seule image est suffisante pour calculer la profondeur 3D des fenêtres. Dick *et al.* introduisent une méthode qui génère des modèles à partir de plusieurs images [34]. Ils présentent une génération statistique d'un modèle basée sur des méthodes d'échantillonnage MCMC (*Markov Chain Monte Carlo*) à partir de distributions de probabilité. La méthode essaye de déduire les propriétés sémantiques des bâtiments afin d'aider la modélisation. Une étendue de cette approche est présentée en [104]. Ils utilisent également l'ISM (*Implicit Shape Model*) pour la détection de fenêtres. L'ISM est la description d'un objet sous la forme de l'organisation spatiale des principales caractéristiques. L'apprentissage d'un modèle de fenêtre a été présenté dans [103]. Des méthodes qui utilisent aussi les données 3D (Lidar terrestre) ont été présentées dans [132, 133]. Cette reconstruction est basée sur la méthode

rjMCMC ( *reversible jump MCMC*) et sur les grammaires de forme. Une autre application des grammaires 2D a été présentée par Alegre [4], où les images de façades des bâtiments modernes sont interprétées en utilisant la combinaison d'un modèle stochastique bayésien de grammaire non contextuelle et un échantillonnage MCMC. D'un autre point de vue, les fenêtres ont aussi été détectées en utilisant des méthodes d'apprentissage [5, 78, 149]. Les caractéristiques de fenêtres sont extraites (ondelettes de Haar) puis une méthode de classification est appliquée : AdaBoost dans [5, 149] et séparateurs à vastes marges (SVM *Support Vector Machine* [160]) dans [78].

## 2.6 Discussion

Dans ce chapitre, nous avons parcouru les différentes techniques qui contribuent à la modélisation d'environnements urbains. Par rapport à nos types de données, nous sommes confrontés à plusieurs défis dans l'extraction de l'information d'intérêt. Par exemple, l'acquisition des données est réalisée dans Paris, sans intervenir sur l'activité naturelle de la ville. Nous rencontrons différents types de véhicules (voitures, bus, motos, etc...), des piétons, ainsi que l'ensemble du mobilier urbain<sup>9</sup> (des lampadaires, des panneaux de signalisation, etc...) qui sont considérés comme des artefacts pour la modélisation de la scène. Ces artefacts doivent être supprimés afin de faciliter la modélisation des façades et du sol. Les techniques existantes d'extraction d'artefacts sur des données 3D ont des contraintes par rapport aux systèmes d'acquisition [56, 97]. Dans notre cas d'étude, des techniques qui dépendent de la géométrie de la scène et non des systèmes devront être développées car nous travaillons avec les données 3D issues de trois systèmes différents. De plus, la classification de ces artefacts sera étudiée parce que le mobilier urbain pourrait être réintroduit dans la scène modélisée, en augmentant ainsi son réalisme. Par ailleurs, dans le cas des images étudiées, les approches [87, 114, 170] travaillant avec des images relativement simples qui ne se généralisent pas dans le cas des façades texturées, ne pourront pas être utilisées. De nouvelles approches, plus robustes, seront donc nécessaires afin d'analyser les façades haussmanniennes car elles sont vraiment riches en détails architecturaux tels que les balcons, les dalles, les jambages de fenêtres, les linteaux, etc. Par ailleurs, lors de l'acquisition, l'ensemble de bâtiments d'une rue est obtenu. Néanmoins, les approches de modélisation s'appliquent à un seul bâtiment [55, 104, 172] ou à une seule façade, découpés manuellement pour extraire des imagerie [72, 114, 133]. Nous devons donc segmenter les données, d'abord en îlots et ensuite en bâtiments, avant d'entamer la modélisation elle-même. Toutes ces problématiques mentionnées sont abordées au cours de cette thèse.

---

9. Sur la ville de Paris, il y a environ 200 types d'éléments de mobilier urbain. Information extraite du *Cahier des Normes d'établissement et d'exploitation des plans de voirie de Paris*.

## **Deuxième partie**

### **Extensions de l'Ouverture Ultime**

# Ouverture Ultime par Attributs

## Sommaire

<b>3.1</b>	<b>Introduction</b>	<b>33</b>
<b>3.2</b>	<b>Opérateurs connexes et zones plates</b>	<b>34</b>
<b>3.3</b>	<b>Ouverture par attributs</b>	<b>34</b>
<b>3.4</b>	<b>Ouverture ultime</b>	<b>35</b>
<b>3.5</b>	<b><i>Max-Tree</i></b>	<b>37</b>
<b>3.6</b>	<b>Ouverture ultime sur des images réelles</b>	<b>40</b>
<b>3.7</b>	<b>Problèmes de l'ouverture ultime</b>	<b>40</b>
3.7.1	Structures imbriquées	40
3.7.2	Transitions graduelles	42
3.7.3	Problèmes de l'ouverture ultime sur des images réelles	44
3.7.4	Stratégies pour résoudre les cas problématiques de l'opérateur	44
<b>3.8</b>	<b>Discussion</b>	<b>46</b>

## 3.1 Introduction

Dans ce chapitre nous abordons les définitions de morphologie mathématique en relation avec l'ouverture ultime par attributs. Premièrement, nous introduisons les conventions utilisées dans ce travail. Puis, les définitions d'un opérateur connexe et des zones plates sont présentées. Ensuite, nous décrivons l'ouverture ultime et son implémentation en utilisant une approche *Max-Tree*. Les cas problématiques de l'opérateur sont illustrées dans la section 3.7. Pour finir, nous discutons des points forts et des points faibles de cet opérateur dans la dernière section.

Une image binaire  $X$  est définie comme un sous-ensemble du domaine  $E \subseteq \mathbb{R}^n$  ou  $\mathbb{Z}^n$  (généralement  $n = 2$ ). Une image à niveaux de gris  $I$  est définie comme un sous-ensemble  $E \rightarrow \mathcal{T}$  tel que  $\mathcal{T} = \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  dans le cas des images continues ou  $\mathcal{T} = \overline{\mathbb{Z}}$  dans le cas des images à valeurs discrètes.  $\mathcal{C}(E)$  est l'ensemble des sous-ensembles connexes de  $E$  et  $C$  un ensemble quelconque en  $\mathcal{C}(E)$  tel que,  $C \in \mathcal{C}(E)$  [146]. La notation  $\bigvee A$

désigne le supremum de l'ensemble  $A$  et  $\bigwedge A$  l'infimum.

## 3.2 Opérateurs connexes et zones plates

La définition d'opérateur connexe a été formalisée par Serra et Salembier dans [140]. Cet opérateur est la généralisation des ouvertures par reconstruction présentées par Serra dans [147].

Soit une différence symétrique de deux ensembles  $A$  et  $B$ , l'ensemble étant constitué par la réunion des éléments de  $A$  qui ne sont pas dans  $B$ , et des éléments de  $B$  qui ne sont pas dans  $A$  :  $(A \Delta B) = (A \cap B^c) \cup (A^c \cap B)$ .

**Définition 3.** Un opérateur binaire  $\Psi$  est dit connexe si pour tout ensemble  $A \in E$ ,  $A \Delta \Psi(A)$  est exclusivement constitué de sous-ensembles connexes de  $A$  ou de son complémentaire  $A^c$  [140] :

$$\Psi \text{ est connexe si } \mathcal{C}(A \Delta \Psi(A)) \subset \mathcal{C}(A) \cup \mathcal{C}(A^c) \quad (3.1)$$

La Définition 3 montre que l'opérateur consiste à préserver ou à supprimer des composants connexes. Le passage de l'opérateur  $\Psi$  aux images à niveaux de gris est fait en utilisant la définition des zones plates (*Flat Zones*,  $FZ$ ). Une zone plate d'une image  $I$ , notée  $L_h(I)$ , est définie comme la région connexe la plus grande possible de niveau de gris constant  $h$  (pas de voisins de valeur  $h$ ). Les ensembles de zones plates de niveau  $h$  sont définis comme  $\mathcal{W}_h(I) = \{x \in E | I(x) = h\}$  et appelés une partition en zones plates par toutes les valeurs de  $h \in I$ . Alors, un opérateur connexe  $\Psi$  à niveaux de gris est appliqué au niveau des zones plates.

**Définition 4.** Un opérateur  $\Psi$  à niveaux de gris est dit connexe si la partition en zones plates de  $\Psi(I)$  est moins fine que celle de  $I$ .

$$\forall FZ(I) \subseteq FZ(\Psi(I)) \quad (3.2)$$

Ces opérateurs permettent la simplification en conservant et/ou en supprimant des zones plates, et donc, ils ne modifient pas la valeur individuelle du pixel. De plus, ils ne peuvent pas créer de nouveaux contours ni changer la position de contours déjà existants, c'est-à-dire n'introduisent aucune nouvelle transition de niveaux de gris dans les images. L'implémentation efficace d'algorithmes à base de files d'attente [164], *Max-tree* [139] et *union find* [174], ont largement contribué à la diffusion de ces opérateurs. Une comparaison des trois méthodes a été présentée dans [105].

## 3.3 Ouverture par attributs

Les ouvertures par attributs ont été formalisées par Breen et Jones dans [20]. Elles consistent en des ouvertures connexes associées à un critère croissant  $T$ , fonction d'un ou plusieurs attributs  $\kappa$ . Ce critère est utilisé pour garder ou enlever des composantes connexes. Voyons d'abord le cas binaire :

**Définition 5.** Soit  $X \subseteq E$  un ensemble et  $T$  un critère croissant. Une ouverture par attribut  $\Gamma^T$  est décrite par :

$$\Gamma^T(X) = \bigcup_{x \in X} \Gamma_T(\Gamma_x(X)) \quad (3.3)$$

Où,  $\Gamma_x$  est l'ouverture connexe de l'ensemble  $X$  à un point  $x$ ,  $\Gamma_T$  est l'ouverture triviale pour la vérification du critère  $T$ .

L'ouverture binaire peut être étendue à des niveaux de gris en appliquant l'ouverture à chaque seuil de l'image et en superposant les résultats.

**Définition 6.** Soit  $I$  une image à niveaux de gris et  $T$  un critère croissant. Une ouverture par attribut  $\gamma^T$  est décrite par :

$$\gamma^T(I(x)) = \max(h \mid x \in \Gamma^T(\mathcal{J}_h(I))) \quad (3.4)$$

Où,  $\mathcal{J}_h(I) = \{x \in E \mid I(x) \geq h\}$  est un seuillage à niveau  $h$ , et  $\Gamma^T$  est une ouverture binaire par critère.

Le critère assigne une valeur booléenne qui indique si une relation est vérifiée ou non. Cette relation est fonction d'un ou plusieurs attributs d'un ensemble connexe  $C \subseteq E$ . Par exemple, Vincent a introduit l'attribut de surface [164] et Urbach *et al.* ont proposé un critère à partir d'un vecteur d'attributs [159].

**Définition 7.** Soit  $C$  un sous-ensemble connexe de  $E$  et  $T$  un critère associé à un attribut  $\kappa$ .  $T(C) = \rho(\kappa_C)$  où  $\rho : \mathbb{R} \rightarrow \{faux, vrai\} \forall C \subseteq E$ . Si  $T(C)$  est vrai on dit que  $C$  vérifie le critère  $T$ .

Par ailleurs, comme nous l'avons déjà mentionné, le critère utilisé doit être croissant pour garantir la définition d'une ouverture. Si le critère est non-croissant, on parlera alors d'un amincissement.

**Définition 8.** Un critère  $T$  est croissant, si le fait qu'un ensemble connexe  $C$  vérifie le critère  $T$  implique que tous les ensembles  $B \supseteq C$  vérifient le critère  $T$ .

Un des critères croissants le plus répandu est la relation  $T(C) : \kappa_C \geq \lambda$ , où  $\kappa_C$  est un attribut croissant de la composante connexe  $C$  et  $\lambda$  un paramètre donné. En utilisant ce type de critères, une ouverture par attribut  $\gamma^T$  peut être dénotée comme  $\gamma_\lambda$ . La Figure 3.1 illustre plusieurs attributs croissants. Ces attributs dépendent normalement de deux types de configuration : le plus petit polygone circonscrit à la composante connexe et le plus grand polygone inscrit dans la composante connexe.

### 3.4 Ouverture ultime

L'ouverture ultime  $\theta$  est un opérateur résiduel qui analyse la différence entre des ouvertures consécutives d'une même famille. Cet opérateur morphologique peut être utilisé pour l'analyse de formes par l'association d'une

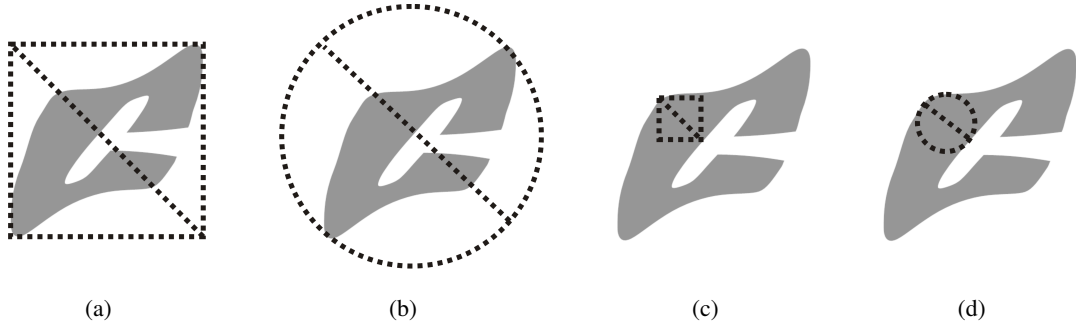


FIGURE 3.1 – (a) Les mesures du plus petit rectangle (boîte englobante) qui contient  $C$  telles que : la surface, la diagonale, la hauteur (le diamètre de Ferret vertical), la largeur (le diamètre de Ferret horizontal), l'élongation, etc. (b) Les mesures de la plus petite circonférence qui enferme  $C$  telles que : la surface, le rayon, le diamètre (le diamètre maximum de Ferret), etc. (c) Les mesures du rectangle le plus grand contenu dans  $C$ . (d) Les mesures de la circonférence la plus grande contenue dans  $C$ .

fonction de granulométrie. Ses avantages sont multiples : il est non paramétrique, il extrait les structures les plus contrastées et il est intrinsèquement multi-échelle. Cet opérateur possède deux images de sortie associées à une image d'entrée  $I$  : la différence maximale entre ouvertures (Résidu  $R_\theta(I)$ ) et la taille de l'ouverture quand le résidu maximal est généré (Indicatrice  $q_\theta(I)$ ).

**Définition 9.** Soit  $I$  une image à niveaux de gris, une ouverture ultime  $\theta$  est décrite par :

$$\theta(I) : I \xrightarrow{\theta} (R_\theta(I), q_\theta(I)) \quad (3.5)$$

$$R_\theta(I) = \bigvee (r_\lambda(I)), \quad \forall \lambda \geq 1 \quad (3.6)$$

où  $r_\lambda(I) = \gamma_\lambda(I) - \gamma_{\lambda+1}(I)$

$$q_\theta(I) = \bigvee (\lambda + 1 : \lambda \geq 1, \quad \text{si } \begin{cases} R_\theta(I) = r_\lambda(I) \\ \text{et} \\ R_\theta(I) > 0 \end{cases} \quad (3.7)$$

où,  $\gamma_\lambda(I)$  est une ouverture de taille  $\lambda$ .

Diverses applications ont été développées en utilisant l'ouverture ultime : l'analyse d'images pour mesurer la granulométrie de roches [118], la localisation automatique de texte enfoui [129], la binarisation d'images de documents [64].

Dans cette thèse, nous utiliserons l'approche *Max-Tree* pour la mise en œuvre de l'ouverture ultime par attributs et pour la présentation des améliorations apportées.



### 3.5 Max-Tree

Le *Max-Tree*, par dualité *Min-Tree*, a été introduit par Salembier [139] comme une structure pour calculer des opérateurs connexes. Il s'agit d'une représentation multi-échelle d'une image et d'une structure hiérarchique en forme d'arbre de composantes connexes. Les nœuds de l'arbre,  $C_h^k$ , représentent les  $k$  composantes connexes binaires de  $\mathcal{J}_h(I)$ . La racine correspond à toute l'image. Les feuilles correspondent aux maxima de l'image. Les liens entre les nœuds (branches) décrivent la relation d'inclusion entre composantes connexes binaires. Une fois que le *Max-Tree* est créé, plusieurs attributs peuvent être calculés, permettant la mise en œuvre efficace des ouvertures par attributs. Le lecteur pourra consulter [116, 139, 173] pour la description des différents algorithmes de création du *Max-Tree*.

La Figure 3.2 illustre une image synthétique (Figure 3.2(a)) et sa correspondante représentation *Max-Tree*. L'ensemble de seuils nécessaires pour construire l'arbre sont présentés dans la Figure 3.2(b). Le *Max-Tree* est une simplification de l'arbre de composantes connexes classique<sup>1</sup>. Comme nous pouvons le constater, le *Max-Tree* a deux feuilles, ce qui correspond au nombre de maxima de l'image.

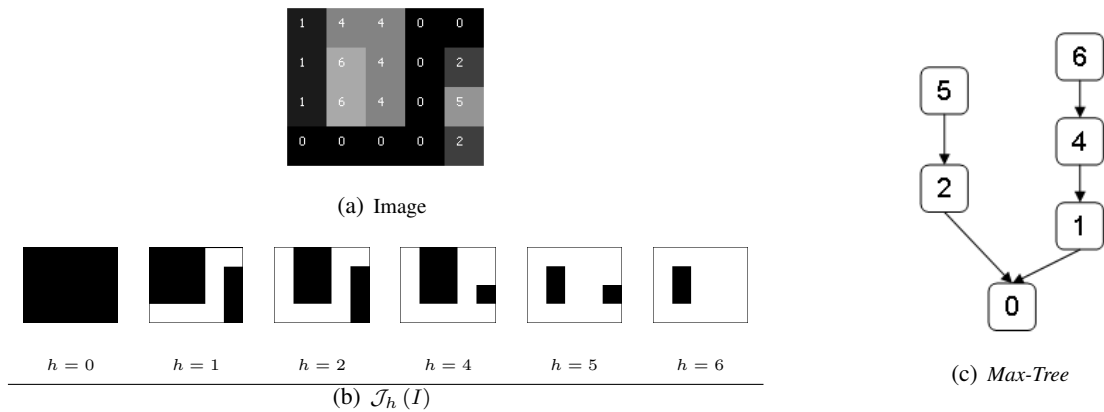


FIGURE 3.2 – (a) Image Synthétique, (b)  $\mathcal{J}_h(I) \forall h \in E$ , (c) *Max-Tree*

La mise en œuvre d'un filtre connexe basé sur un *Max-Tree* consiste en l'élagage de l'arbre suivi d'une restitution de l'image. Dans le cas particulier des opérateurs par attributs, l'élagage de branches se fait suivant une stratégie d'élimination de nœuds en fonction d'un critère  $T$ . Salembier a décrit quatre différentes stratégies : Min, Max, Directe et Viterbi pour supprimer des nœuds quand le critère est non-croissant [139]. Pour les ouvertures par attributs  $\gamma^T$  (critère croissant) un nœud est éliminé s'il ne vérifie pas le critère  $T$ . Avec la supposition de  $T(A) : \kappa_A \geq \lambda$ , l'ouverture par attribut  $\gamma_\lambda$  supprime un nœud  $C_h^k$  quand son attribut  $\kappa_{C_h^k}$  est plus petit qu'un paramètre donné  $\lambda$ .

1. L'arbre de composantes connexes classique *ACCC* présente un nœud pour chaque zone plate de l'image. Dans l'exemple présenté le *ACCC* a 10 nœuds alors que le *Max-Tree* n'en a que 6.

Un exemple d'ouverture par attribut de hauteur est illustré sur la Figure 3.3. L'arbre illustre par des traits pointillés les branches élaguées pour les différentes valeurs d'attribut  $\lambda$ . Tout d'abord, l'ouverture  $\gamma_2$  élimine le nœud  $C_5^0$ , ensuite le nœud  $C_6^0$  est supprimé par l'ouverture  $\gamma_3$ . Notez que les nœuds  $C_1^0$  et  $C_4^0$  de la même branche et le nœud  $C_2^0$  ont une hauteur  $\kappa = 3$  et ils sont donc supprimés par l'ouverture  $\gamma_4$ .

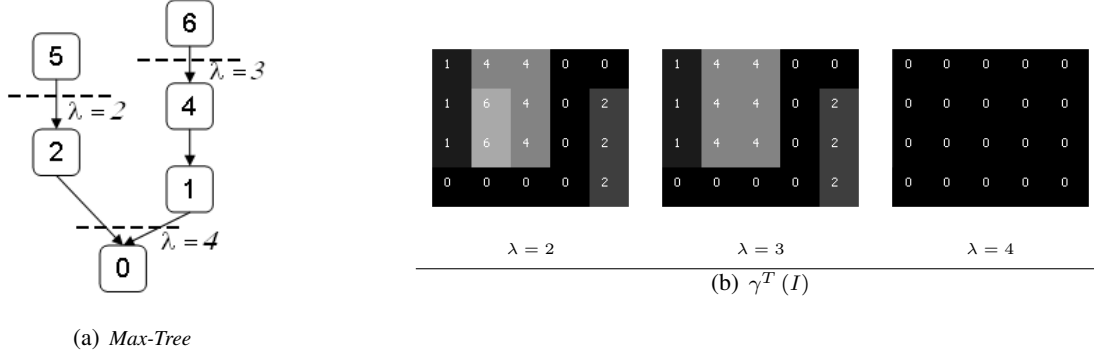


FIGURE 3.3 – (a) *Max-Tree* et trait pointillé aux branches à élaguer (b) ouverture par hauteur  $\gamma_T(I)$ , où  $T : \kappa_{C_h^k} \geq \lambda$ .

L'ouverture ultime par attributs peut être calculée sur le *Max-Tree*, de la même manière que l'ouverture par attributs, car elle est un opérateur connexe. Fabrizio a décrit la mise en œuvre en utilisant la structure du *Max-Tree* [41]. Le résidu  $r_\lambda$  de chaque nœud supprimé par une ouverture  $\gamma_\lambda$  est calculé par la différence entre son niveau de gris  $h$  et le niveau de gris de son premier ancêtre avec un attribut différent.

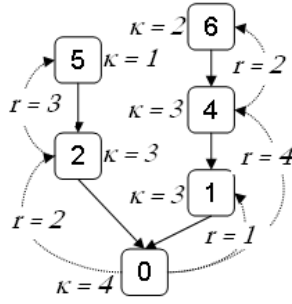
La Figure 3.4 montre le calcul des résidus sur chaque nœud du *Max-Tree* et les images intermédiaires utilisées pour le calcul de l'ouverture ultime par attribut de hauteur ( $R_\theta(I)$ ,  $q_\theta(I)$ ). Analysons l'arbre de manière séquentielle<sup>2</sup>, de la racine vers les feuilles, tel que cela a été décrit dans [41]. Rappelons que la région correspondant au nœud père englobe l'ensemble des régions des nœuds enfants. C'est pourquoi, une fois un nœud parent supprimé par une ouverture donnée ( $\kappa > \lambda_i$ ), ses enfants sont automatiquement éliminés et deviennent indivisibles pour des ouvertures  $\kappa > \lambda_j > \lambda_i$ . De cette manière, la suppression d'un nœud, pour obtenir une ouverture donnée, génère un résidu pour l'ensemble des pixels des régions de tous les nœuds descendants. Pour cette raison l'ordre de parcours de la racine vers les feuilles permet de propager l'information de résidus de manière efficace et de calculer le résidu maximal pour chaque nœud en un seul parcours de l'arbre. Ainsi, le calcul de l'ouverture ultime par hauteur de l'image se fait de la manière suivante :

1. Les nœuds  $C_1^0$ ,  $C_4^0$  et  $C_2^0$  correspondent aux zones plates de l'image qui ont changé entre  $\gamma_3$  et  $\gamma_4$ . La différence  $\gamma_3 - \gamma_4$  produit les résidus suivants :  $r_4(C_2^0) = 2$ ,  $r_4(C_1^0) = 1$  et  $r_4(C_4^0) = 4$ . Comme le nœud  $C_4^0$  a le même attribut que son parent  $C_1^0$ , le résidu est la différence entre son niveau de gris et le niveau de

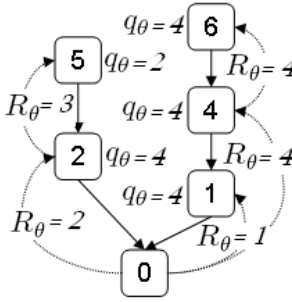
2. Le calcul de l'opérateur de façon séquentielle est fait pour introduire la mise en œuvre sur l'arbre. Néanmoins, les calculs des résidus peuvent être réalisés dans n'importe quel l'ordre.

gris de son parent, plus le résidu de son parent (voir équation 3.8). Ces premiers résidus sont utilisés pour la mise à jour des valeurs de transformée  $R_\theta(C_h^k) = r_4(C_h^k)$  et d'indicatrice  $q_\theta(C_h^k) = 4$ .

2. L'ouverture  $\gamma_3$  coupe la branche entre les nœuds  $C_6^0$  et  $C_4^0$ , produisant un résidu de  $6-4=2$ , ce qui est inférieur au résidu du père ( $=4$ ), c'est pourquoi  $R_\theta(C_6^0) = 4$ .
3. L'ouverture  $\gamma_2$  coupe une dernière branche (entre les nœuds  $C_5^0$  et  $C_2^0$ ), produisant un résidu de  $5-2 = 3$ , supérieur au résidu du père. Il est donc affecté à  $R_\theta(C_5^0) = 3$  et  $q_\theta(C_5^0) = \kappa_{C_5^0} + 1 = 2$ .



(a) Max Tree +  $r_\lambda(C_h^k)$



(c) Max Tree +  $R_\theta(C_h^k), q_\theta(C_h^k)$

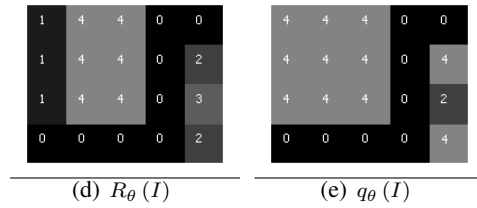
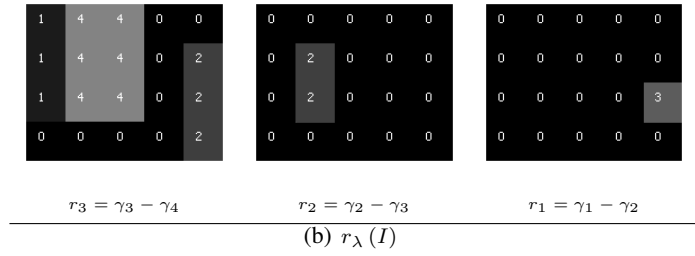


FIGURE 3.4 – (a) Max-Tree et résidu  $r_\lambda(C_h^k) \forall h \in E$ , (b) les images intermédiaires de  $r_\lambda(I)$ , (c) Max-Tree et  $R_\theta(C_h^k), q_\theta(C_h^k)$  et les images résultats de : (d)  $R_\theta(I)$ , et (e)  $q_\theta(I)$  de l'ouverture ultime par hauteur.

Ce calcul de  $r_\lambda$  est résumé dans une procédure itérative de la manière suivante :

$$r_\lambda(C_h^k) = \begin{cases} h - h' + r_\lambda(C_{h'}^{k'}) & \kappa_{C_h^k} = \kappa_{C_{h'}^{k'}} \\ h - h' & \text{sinon} \end{cases} \quad (3.8)$$

où,  $C_{h'}^{k'}$  est le parent du nœud ( $C_h^k \subset C_{h'}^{k'}$  et  $h' < h$ ). Dans un premier temps, le résidu de la racine (premier parent) est initialisé à zéro. Ensuite, les résidus des enfants sont calculés entre le parent et ses enfants. Chaque enfant compare son résidu avec le résidu maximal de son parent  $R_\theta(C_{h'}^{k'})$ , et  $R_\theta(C_h^k)$  enregistre la valeur maximale

entre eux. Ensuite, les enfants deviennent parents et répètent le processus.  $q_\theta(C_h^k)$  est la valeur de l'attribut de l'enfant plus un, quand le résidu maximal est généré.

### 3.6 Ouverture ultime sur des images réelles

La Figure 3.5 présente trois exemples de l'utilisation de l'ouverture ultime par attribut sur des images réelles. Les figures représentent les types d'image traités dans les applications qui seront étudiées dans le Chapitre 7.

Il faut rappeler que les résultats des trois images ont été obtenus avec le même opérateur non-paramétrique, ce qui montre la puissance de l'approche. Un premier exemple de segmentation de façades illustre la puissance de l'opérateur pour l'extraction de structures internes comme les fenêtres. Dans cet exemple l'opérateur a été utilisé sur l'inverse d'une image de gradient couleur (voir Figure 3.5(a)). Un autre exemple présente l'utilisation de l'opérateur dans la segmentation de texte enfoui. L'opérateur présente de bons résultats de segmentation des lettres même si celles-ci n'ont pas de couleurs homogènes à l'intérieur (voir Figure 3.5(b)). La dernière image illustre une application de segmentation des images de cellules. Les cellules sont bien segmentées sauf une qui a été sur-segmentée.

### 3.7 Problèmes de l'ouverture ultime

L'ouverture ultime présente plusieurs avantages comme méthode de segmentation : elle est non-paramétrique, intrinsèquement multi-échelle et possède la propriété d'extraire les structures les plus contrastées. Néanmoins, ce dernier avantage peut devenir son plus grand inconvénient. Rappelons la définition de l'ouverture ultime : l'opérateur enregistre le dernier résidu maximal calculé. Il ne garde pas d'informations intermédiaires, qui pourraient être aussi importantes. Retornaz, dans sa thèse, a décrit ces types de problèmes dus à l'absence de mémoire de l'opérateur [129]. Ces problèmes sont principalement produits par deux types de configuration : des structures imbriquées et des transitions graduelles. Par exemple, une configuration pyramidale dans l'image est un cas classique de structures imbriquées.

#### 3.7.1 Structures imbriquées

Ce type de problème est mis en évidence quand les structures d'intérêt se trouvent imbriquées sur des composantes de plus fort contraste. Par conséquent, quand une structure est contenue dans une autre, elle peut être masquée par un plus grand résidu. Pour mieux comprendre ce problème, analysons les images de la Figure 3.6.

L'image synthétique de la Figure 3.6(a) a trois composantes imbriquées : un rectangle (dimensions  $120 \times 40$ ,  $h = 250$ ), un carré (dimensions  $30 \times 30$ ,  $h = 200$ ) et un cercle (diamètre 90,  $h = 175$ ). Le carré se trouve également contenu dans le cercle. Les trois structures se trouvent sur un plateau carré de dimensions  $160 \times 160$

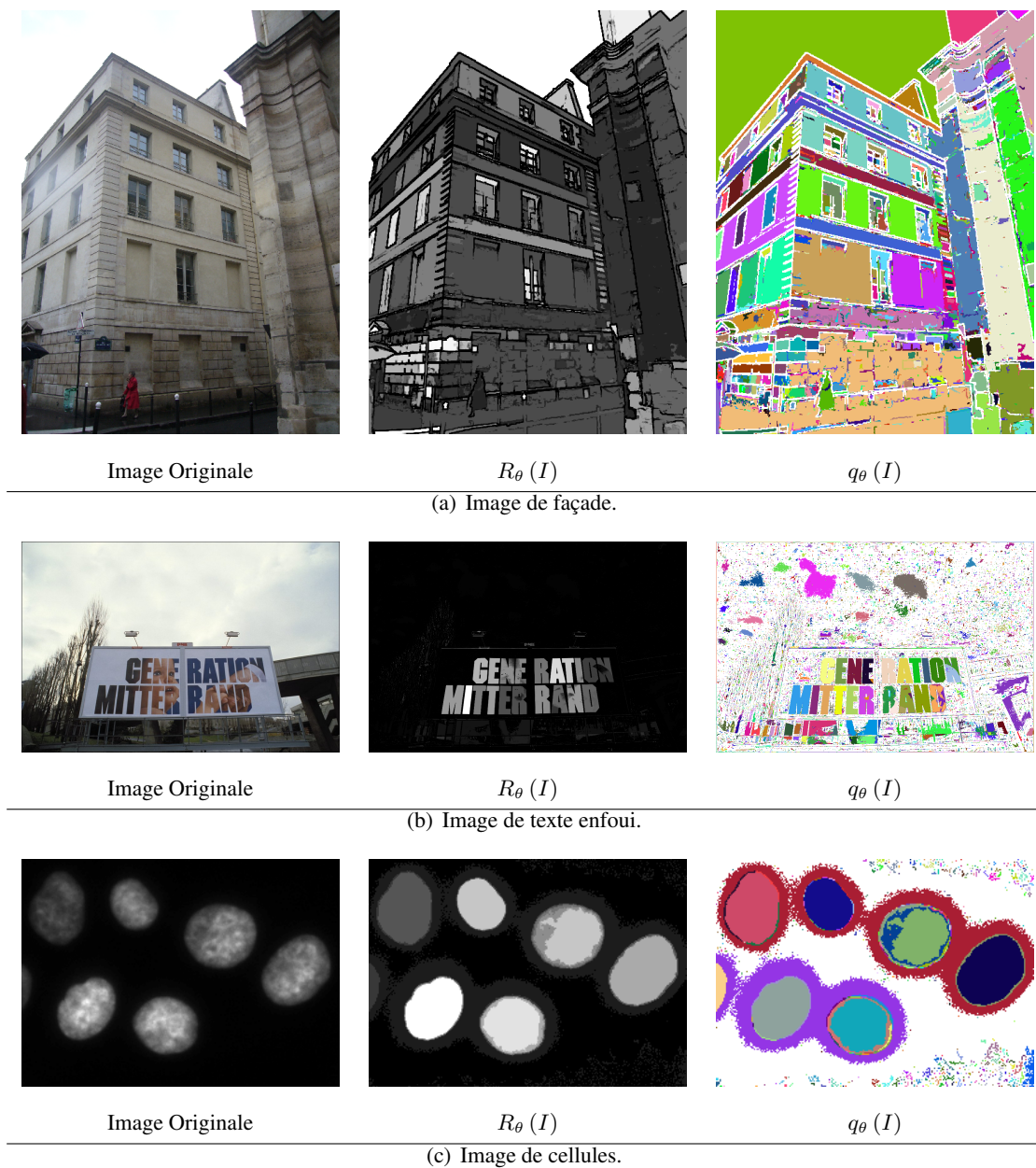


FIGURE 3.5 – Ouverture ultime sur des images réelles.

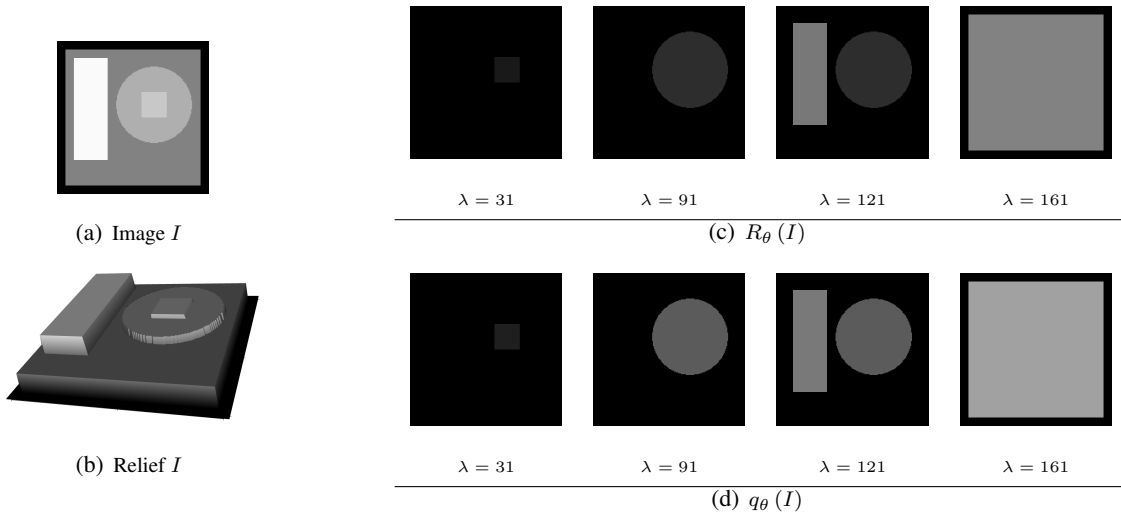


FIGURE 3.6 – Exemple : Problème dû aux structures imbriquées (a) Image Synthétique, (b) Représentation en relief d' $I$ , (c) Séries de transformées et (d) Séries d'indicatrices.

et au niveau  $h = 130$ . Si on applique la définition de l'ouverture ultime par hauteur pas à pas, on obtient le comportement suivant :

1. Le carré est la première structure rencontrée suite à une ouverture de taille  $\lambda = 31$  avec un résidu égal à  $200 - 175 = 25$ .
2. Le carré est masqué par le cercle après l'ouverture de taille  $\lambda = 91$  avec un résidu égal à  $175 - 130 = 45$ .
3. Le rectangle est détecté avec l'ouverture de taille  $\lambda = 91$  et le résidu égal à  $250 - 130 = 120$ .
4. Les trois structures sont masquées pour le résidu du plateau avec le bord à 0 égal à 130.

### 3.7.2 Transitions graduelles

Les transitions graduelles des structures produisent dans l'opérateur des petits résidus. C'est pour cette raison que les structures qui présentent des transitions graduelles qui se trouvent imbriquées ont de fortes chances d'être masquées. Un exemple de ce problème est illustré dans la Figure 3.7. L'image synthétique de la Figure 3.7(a) est composée d'un rectangle de dimensions  $120 \times 40$  et au niveau  $h = 250$ . Le rectangle a une transition dans la partie inférieure des dimensions  $2 \times 40$  au niveau  $h = 170$  (la transition peut être facilement observée dans la Figure 3.7(b)). Le rectangle est placé sur un plateau carré de dimensions  $160 \times 160$  et niveau  $h = 110$ . En utilisant l'ouverture ultime par hauteur, le rectangle est masqué par le plateau, parce qu'il ne voit que des résidus de  $250 - 170 = 80$  et  $170 - 110 = 60$  au lieu de  $250 - 110 = 140$  (La valeur de 140 est la différence des niveaux de gris du rectangle et du plateau dans le cas où la transition n'existait pas). Le résidu produit par le plateau est égal à 110 et donc plus grand que les résidus du rectangle ( $110 > 80$ ).

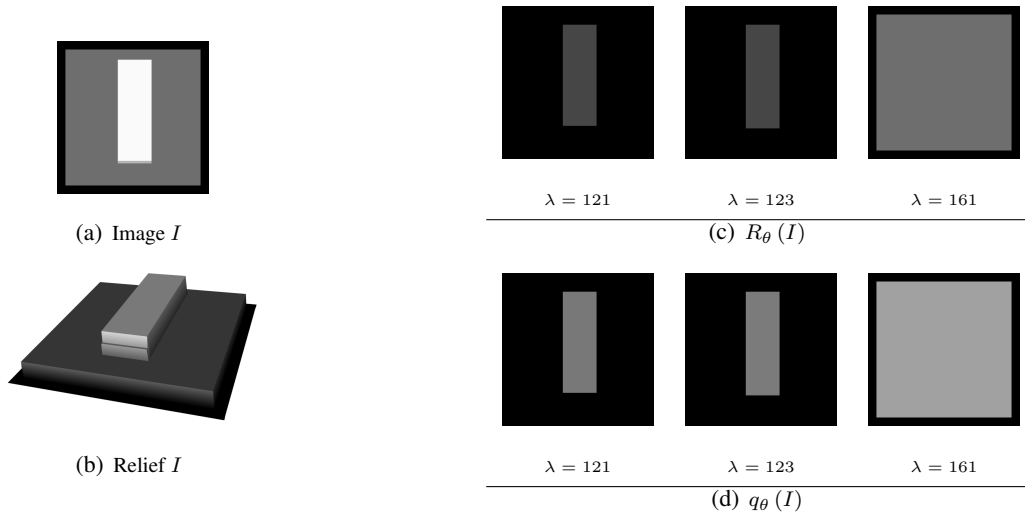


FIGURE 3.7 – Exemple : Problème dû aux transitions graduelles (a) Image Synthétique, (b) Représentation en relief d' $I$ , (c) Séries de transformées et (d) Séries d'indicatrices.

Par ailleurs, ce problème de transitions graduelles peut également générer un problème de fuites quand il s'agit d'ouvertures connexes. Ce problème a été illustré par Salembier dans [138] comme un problème classique des opérateurs connexes. Il est produit par d'indésirables ponts étroits, reliant des composantes séparées de l'image.

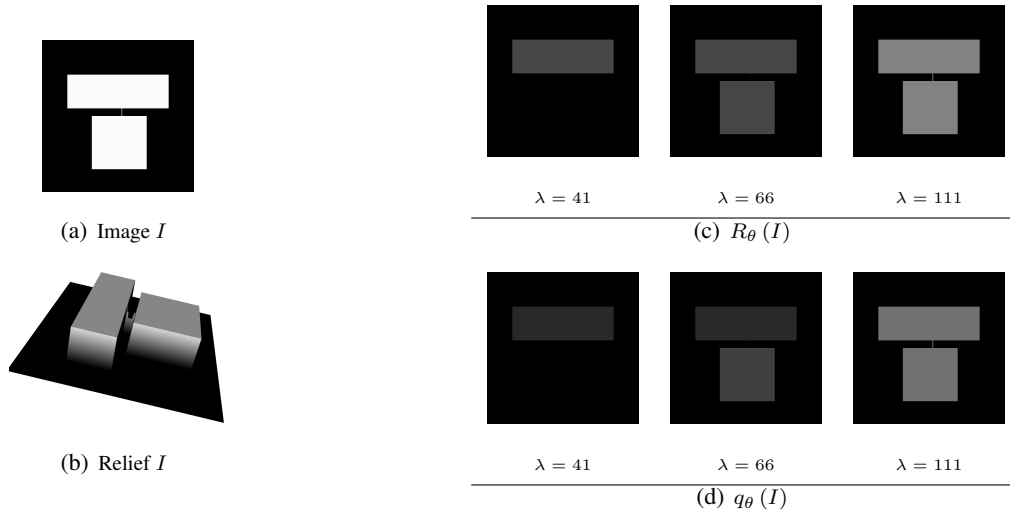


FIGURE 3.8 – Exemple : Problème dû aux problèmes de fuites (a) Image Synthétique, (b) Représentation en relief d' $I$ , (c) Séries de transformées et (d) Séries d'indicatrices.

La Figure 3.8 illustre le problème de fuite en utilisant l'ouverture ultime par hauteur. L'image synthétique de la Figure 3.8(a) est composée de deux structures : un rectangle (dimensions  $40 \times 120$ ,  $h = 250$ ) et un carré (dimensions  $65 \times 65$ ,  $h = 250$ ). Les structures sont liées par 5 pixels de niveau  $h = 180$ . Le résultat final de

l'opérateur est une seule composante, car le résidu généré par la fusion de deux composantes ( $r_{113} = 180$ ) est plus grand que les résidus de chacune des structures séparément ( $r_{41} = r_{66} = 250 - 180 = 70$ ).

### 3.7.3 Problèmes de l'ouverture ultime sur des images réelles

La Figure 3.9 illustre des exemples des problèmes que nous venons d'évoquer sur des images réelles dans les trois applications déjà mentionnées. Dans la Figure 3.9(a), les structures internes de la façade ont été masquées. Il s'agit d'un cas de structures imbriquées parce que le contraste entre le mur et la fenêtre est plus petit que le contraste entre le ciel et la façade du bâtiment. Un exemple de transitions graduelles est présenté dans la Figure 3.9(b). Comme on peut le constater, toutes les lettres sont masquées parce qu'elles ont des bords flous causés par le processus de numérisation. La Figure 3.9(c) montre la fusion de cellules causée par les problèmes de fuites. L'image montre deux cellules légèrement séparées ; cependant en utilisant l'opérateur, l'image résultat présente les deux cellules fusionnées.

### 3.7.4 Stratégies pour résoudre les cas problématiques de l'opérateur

Retornaz [129] a également proposé des alternatives pour combattre les problèmes de l'opérateur. Concernant les structures imbriquées, il a étudié deux stratégies : la taille d'arrêt et les hiérarchies de résidus. Pour le cas des transitions graduelles trois stratégies ont été proposées : le changement du pas de l'ouverture, le redressement des frontières et l'accumulation de résidus.

1. **Taille d'arrêt** : La taille d'arrêt est la taille de la dernière ouverture utilisée dans le processus. Par exemple dans le cas de la Figure 3.6, une taille d'arrêt égale à 160 éviterait le masquage du rectangle et du cercle. Néanmoins, le carré resterait masqué. En outre, dans le cas de la Figure 3.8 cette valeur devrait être inférieure à 110 et supérieure à 65 pour garantir la détection des structures. Alors, la sélection de la taille dépendra de l'application ou d'une connaissance *a priori* des structures de l'image.
2. **Hiérarchies de résidus** : L'idée principale de cette approche est d'analyser les  $N$  sous-résidus maximaux au lieu d'un seul. Néanmoins, la sélection de la valeur optimale de  $N$  reste un problème non résolu<sup>3</sup>. Par exemple, dans l'image de la Figure 3.6, les structures se trouvent dans deux niveaux de hiérarchie différents. Alors, si nous analysons le premier niveau hiérarchique, le rectangle et le cercle seront détectés ; et si nous analysons le deuxième niveau, seulement le carré sera détecté.
3. **Changement du pas de l'ouverture** : Dans la définition de l'ouverture ultime, le calcul de résidus est fait à partir d'une série d'ouvertures consécutives. La proposition implique le calcul de l'opérateur avec un pas donné différent de 1. Cependant, sur les exemples présentés, plusieurs pas devront être utilisés afin de

3. Une étude similaire a été présenté par Leite et Guimaraes dans [88]



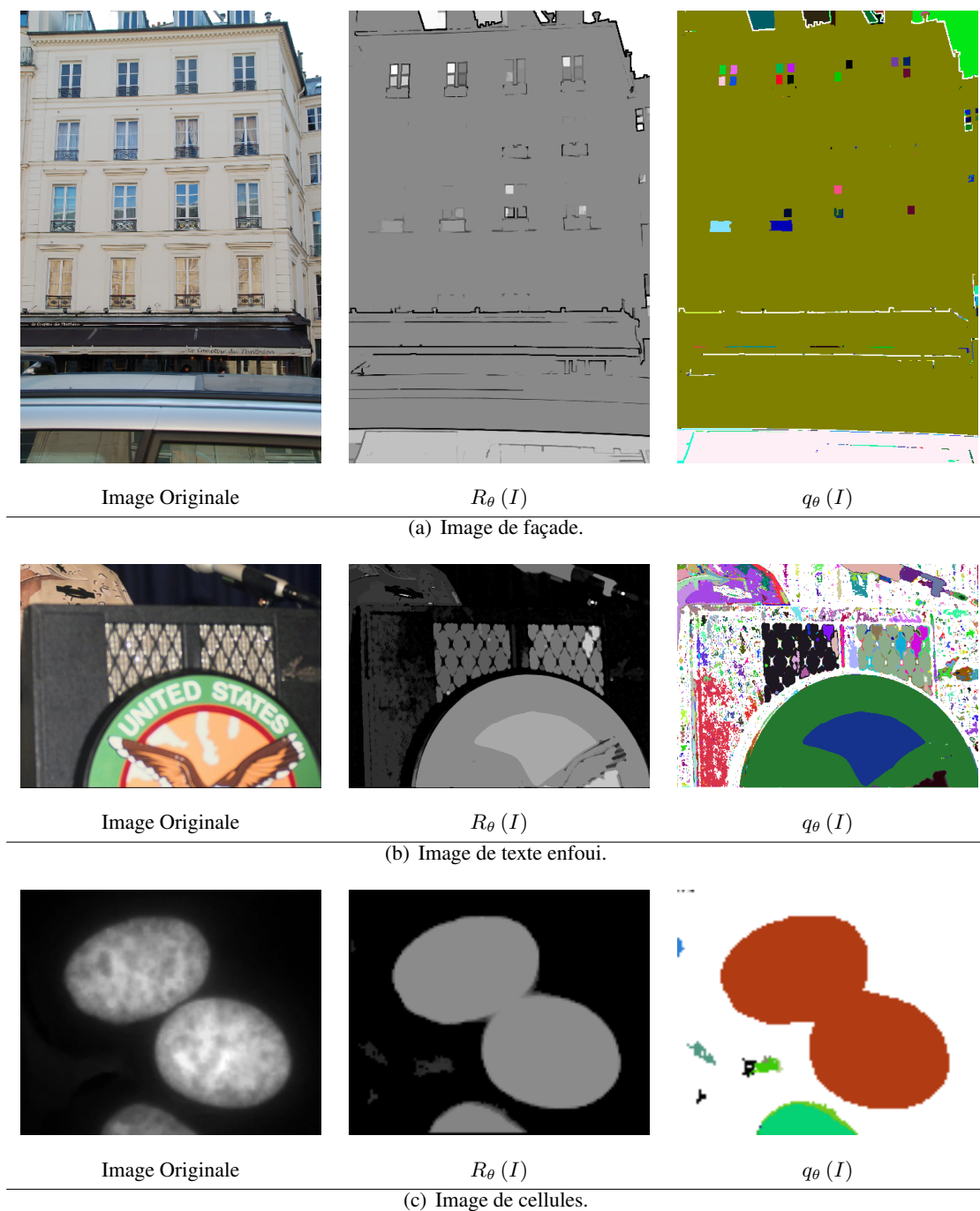


FIGURE 3.9 – Problèmes de l'ouverture ultime sur des images réelles.

résoudre les problèmes de structures imbriquées et de fuites, en détectant les composantes d'intérêt. Ainsi, la sélection d'un pas est difficile de manière générale.

4. **Redressement des frontières** : Cette solution travaille au niveau de l'image d'entrée et non sur l'opérateur lui-même. Cependant, il faudra faire attention à ne pas créer de nouvelles structures qui n'existaient pas avant.
5. **Accumulation de résidus** : L'idée de cette proposition est basée sur l'accumulation de résidus quand la variation de l'attribut est inférieure ou égale à un pas  $\Delta$ . Cette stratégie présente une vraie solution au problème de transitions graduelles. Elle sera étudiée plus largement dans cette thèse au Chapitre 6.

Dans notre travail nous étudierons trois stratégies pour combattre les problèmes et ainsi améliorer les résultats de l'opérateur décrit. La première proposition utilise l'information *a priori* des structures d'intérêt. La méthode favorise les résidus avec une fonction de similarité des structures d'intérêt qui ont une forme déterminée. Cette approche sera présentée au Chapitre 4. La deuxième proposition est une généralisation de la proposition de hiérarchie. Nous analysons plusieurs critères pour définir le niveau de hiérarchie et l'analyse des structures qui se trouvent aux différents niveaux. Le Chapitre 5 décrit cette approche. La dernière proposition est une formalisation de l'accumulation de résidus. Nous l'appellerons ouverture ultime par  $\Delta$  – attributs. La méthode sera présentée au Chapitre 6.

### 3.8 Discussion

Nous avons vu, dans ce chapitre, les définitions des opérateurs connexes et spécialement les ouvertures par attributs (critère croissant). De la même manière, nous avons illustré l'ouverture ultime  $\theta$  et le cas particulier de l'ouverture ultime par attributs. L'opérateur présente deux images de sortie, le résidu maximal ( $R_\theta(I)$ ) et la taille de l'ouverture du résidu maximal ( $q_\theta(I)$ ). L'opérateur n'est pas adapté aux applications de segmentation de régions texturées. Des exemples sur des images réelles présentent la performance de l'opérateur pour la détection des régions homogènes<sup>4</sup>. Nous avons aussi analysé la mise en œuvre de l'opérateur en utilisant de la représentation de l'image sur le *Max-Tree*. Nous avons également illustré ses avantages et ses inconvénients. Parmi ses avantages, nous pouvons mentionner : l'opérateur est non-paramétrique, il est intrinsèquement multi-échelle et il extrait les structures les plus contrastées. Cependant, ce dernier avantage est son plus grand inconvénient, car le plus grand résidu peut produire un masquage ou des liens indiscernables entre les structures d'intérêt. Nous avons identifié trois catégories de cas problématiques de l'opérateur : les structures imbriquées, les transitions graduelles et les problèmes de fuites (quand il s'agit d'une ouverture connexe). Par ailleurs, nous avons présenté les améliorations de l'opérateur décrites dans la littérature et celles proposées dans cette thèse qui seront développées dans les chapitres suivants.

---

4. Une définition de la segmentation de régions homogènes (particules) et segmentation de régions texturées est décrite dans [165].

# Ouverture Ultime avec l'Information de Forme

## Sommaire

4.1	Introduction . . . . .	47
4.2	Similarité entre formes . . . . .	48
4.3	Ouverture ultime avec l'information de forme . . . . .	49
4.4	Exemple d'une image synthétique . . . . .	51
4.5	Application : Images de façade . . . . .	53
4.6	Discussion . . . . .	58

## 4.1 Introduction

Le chapitre précédent a illustré l'opérateur de l'ouverture ultime avec ses avantages et ses inconvénients. Dans ce chapitre nous présentons une première stratégie pour détecter des structures qui restent masquées après l'utilisation de l'opérateur classique. Dans plusieurs applications de segmentation d'images, l'utilisateur connaît de manière *a priori* les structures qu'il cherche à détecter. Et c'est exactement à partir de cette notion qu'est née notre approche. L'idée est de caractériser la forme des objets d'intérêt, établir une fonction de similarité avec une forme connue et introduire cette fonction dans le calcul de l'ouverture ultime. De cette manière, nous cherchons à favoriser l'apparition de régions d'une forme donnée, en réduisant leur masquage par d'autres structures plus contrastées.

Le chapitre est organisé de la manière suivante. La section 4.2 introduit la définition de forme et la fonction de similarité. La section 4.3 décrit notre méthode qui combine la fonction de similarité et l'ouverture ultime. Des applications sur une image synthétique et des images de façade sont présentées dans les sections 4.4 et 4.5, respectivement. La dernière section est consacrée à une discussion sur le nouvel opérateur.

## 4.2 Similarité entre formes

La définition d'une forme a été largement étudiée dans la littérature. Deux études des méthodes pour l'analyse des formes peuvent être trouvées dans [93, 161]. Charpiat *et al.* notent une forme quelconque  $\Omega$ , un sous-ensemble régulier borné de  $D$ , et  $\Gamma$  ou  $\partial\Omega$  sa frontière, une courbe régulière de  $E^2$  [26].

Dans notre contexte, nous nous intéressons à la comparaison de deux formes différentes et à leur mesure de similarité. De nombreuses définitions des fonctions de similarité  $\psi()$  entre deux formes  $(\Omega_i, \Omega_j)$  ont été proposées dans la littérature de vision par ordinateur. La plupart des approches qui exploitent l'information de forme *a priori* dans la segmentation, travaillent avec les contours actifs ou dans le cadre d'un modèle déformable. En revanche, nous nous proposons de définir une fonction de similarité de forme par des attributs  $\kappa_\Omega$ , afin de favoriser des formes spécifiques. Nous utilisons les caractéristiques géométriques (hauteur, largeur, etc.) et leurs relations (relation de remplissage, circularité, moments, etc.).

Afin d'introduire une mesure de similarité entre formes dans l'ouverture ultime, une fonction doit être définie. Cette fonction doit être maximale quand les attributs choisis pour les formes analysées sont égaux. Voyons la Définition 10 :

**Définition 10** (Fonction de similarité). *Nous dénotons une fonction  $\psi_\kappa : \Omega \times \Omega \rightarrow [0, 1] \in \mathbb{R}$  comme une mesure de similarité entre formes qui vérifie pour un attribut  $\kappa$  de deux formes,  $\Omega_i$  et  $\Omega_j$ , les conditions suivantes<sup>a</sup> :*

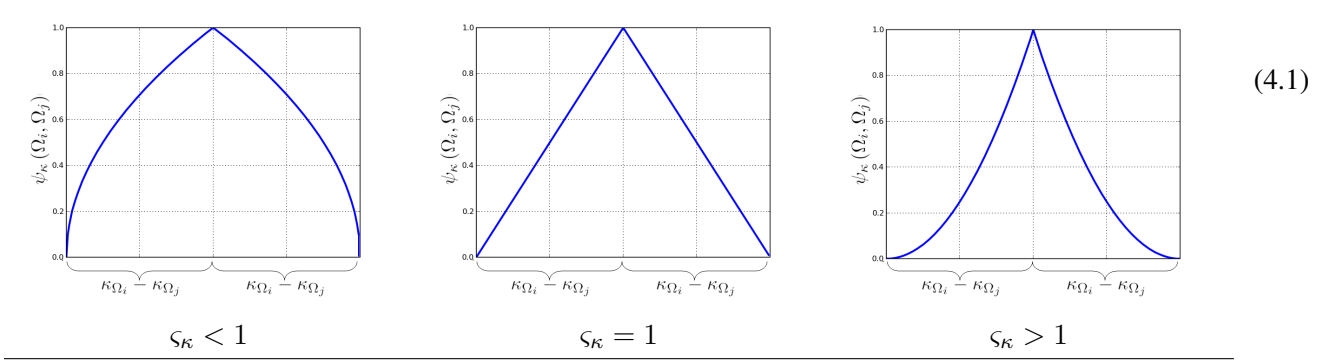
- *Identité* :  $\psi_\kappa(\Omega_i, \Omega_i) = 1$ .
- *Unicité* :  $\psi_\kappa(\Omega_i, \Omega_j) = 1$  implique  $\kappa_{\Omega_i} = \kappa_{\Omega_j}$ .
- *Symétrie* :  $\psi_\kappa(\Omega_i, \Omega_j) = \psi_\kappa(\Omega_j, \Omega_i)$ .
- *Non négativité* :  $\psi_\kappa(\Omega_i, \Omega_j) \geq 0$ .

<sup>a</sup>. Ces conditions sont analogues aux conditions des mesures de distance entre formes présentées dans [162].

Après avoir défini ces conditions, l'équation 4.1 présente un exemple de la fonction de similarité  $\psi_\kappa(\Omega_i, \Omega_j)$ , où  $\tau_\kappa$  est le seuil de similarité de l'attribut  $\kappa$  et  $\varsigma_\kappa$  est un paramètre qui modifie la réponse de la fonction de la manière suivante :

- $\varsigma_\kappa < 1$  : la similarité diminue plus lentement que la différence entre les attributs (courbe convexe).
- $\varsigma_\kappa = 1$  : la similarité dépend linéairement de la différence entre les attributs.
- $\varsigma_\kappa > 1$  : une faible différence entre les attributs entraîne une chute rapide de la fonction de similarité (courbe concave).

$$\psi_{\kappa}(\Omega_i, \Omega_j) \leftarrow \begin{cases} 1 & \kappa_{\Omega_i} = \kappa_{\Omega_j} \\ \left(1 - \left| \frac{\kappa_{\Omega_i} - \kappa_{\Omega_j}}{\tau_{\kappa}} \right| \right)^{\varsigma_{\kappa}} & |\kappa_{\Omega_i} - \kappa_{\Omega_j}| < \tau_{\kappa} \\ 0 & \text{sinon} \end{cases}$$



Nous avons défini une fonction de similarité avec un seul attribut donné  $\kappa$  ; cependant, dans la pratique, plusieurs mesures peuvent être utilisées pour décrire une forme. Ainsi, nous élargirons la comparaison de deux formes  $\Omega_i, \Omega_j$  avec plusieurs attributs en utilisant une simple multiplication des fonctions de similarité :

$$\psi_{\forall \kappa} = \prod_{\forall \kappa} \psi_{\kappa}(\Omega_i, \Omega_j) \quad (4.2)$$

Cette nouvelle fonction  $\psi_{\forall \kappa}$  vérifie également les conditions énoncées par la Définition 10.

### 4.3 Ouverture ultime avec l'information de forme

Nous présentons deux stratégies dans le but d'utiliser l'information de forme avec l'ouverture ultime. Dans la première, nous proposons l'utilisation des ouvertures par des critères de forme, à l'instar de l'approche d'amin-cissement par un vecteur d'attributs présentée dans [159], en utilisant la représentation *Max-Tree*. Cependant, les critères de forme à partir des vecteurs d'attributs ou d'attributs de forme plus complexes comme la circularité, la compacité, etc. . . sont non croissants. C'est pourquoi les stratégies de propagation (Min, Max, Viterbi, et Directe) décrites par Salembier [139] devront être utilisées. Par ailleurs, un autre problème se présente à l'heure de l'utilisation de ces types d'ouverture car l'ouverture ultime est calculée à partir de la différence entre deux ouvertures consécutives. Les filtres par vecteur d'attributs étant définis dans le domaine  $\mathbb{R}$ , un intervalle et un pas d'analyse en  $\mathbb{R}$  sont nécessaires pour calculer  $r_{\lambda}(I) = \gamma_{\lambda}(I) - \gamma_{\lambda+1}(I)$ . C'est pour cette raison que cette approche ne sera pas approfondie dans cette thèse.

Dans la deuxième stratégie, nous proposons de combiner une fonction du facteur de forme  $f(\Omega, \Omega_{ref})$  à partir d'une forme de référence  $\Omega_{ref}$  dans le calcul des résidus (Eq. 4.3). Comme  $\Omega_{ref}$  est un *a priori* fixe,  $f(\Omega, \Omega_{ref})$  est dénotée comme  $f(\Omega)$ .

$$r_{\lambda}^{\Omega} \leftarrow f(\Omega) r_{\lambda} \quad (4.3)$$

De cette manière, le résidu d'une forme  $\Omega$  qui ressemble à  $\Omega_{ref}$  est artificiellement augmenté, en favorisant sa présence dans le résultat de segmentation, même si la forme  $\Omega$  est imbriquée dans une forme plus contrastée différente de la référence.

La fonction du facteur  $f(\Omega)$  est liée à la fonction de similarité  $\psi_{\forall \kappa}$  de  $\Omega$  de la manière suivante :  $1 + \alpha \psi_{\forall \kappa}$ . La constante 1 est ajoutée afin de passer à la version classique de l'opérateur lorsque la fonction de similarité est égale à 0 ( $r_{\lambda}^{\Omega}$  devient  $r_{\lambda}$ ). Le facteur multiplicatif  $\alpha$  est utilisé pour contrôler l'influence de la fonction  $\psi_{\forall \kappa}$  par rapport aux niveaux de gris. Par conséquent,  $1 + \alpha$  représente la valeur maximale que peut atteindre la fonction  $f(\Omega)$ . Finalement, la fonction  $f(\Omega)$  est enregistrée sur une image  $F_{\theta}^{\Omega}(I)$  quand le résidu maximal ( $R_{\theta}(I)$ ) est généré. Avec cette information, nous modifions l'expression originale de la Définition 9 par la Définition 11.

**Définition 11** (Ouverture ultime avec l'information de forme). *Nous définissons une ouverture ultime avec l'information de forme  $\theta^{\Omega}$  d'une image  $I$  par :*

$$\theta^{\Omega}(I) : I \xrightarrow{\theta^{\Omega}} (R_{\theta}^{\Omega}(I), q_{\theta}^{\Omega}(I), F_{\theta}^{\Omega}(I)) \quad (4.4)$$

$$R_{\theta}^{\Omega}(I) = \bigvee (r_{\lambda}^{\Omega}(I)), \quad \forall \lambda \geq 1$$

$$\text{avec } r_{\lambda}^{\Omega}(I) = f(\Omega) \times r_{\lambda}(I) \quad (4.5)$$

$$\left. \begin{array}{l} q_{\theta}^{\Omega}(I) = \bigvee (\lambda) + 1 : \lambda \geq 1 \\ F_{\theta}^{\Omega}(I) = f(\Omega) \end{array} \right\} \text{ si } \left\{ \begin{array}{l} R_{\theta}^{\Omega}(I) = r_{\lambda}^{\Omega}(I) \\ \text{et} \\ R_{\theta}^{\Omega}(I) > 0 \end{array} \right. \quad (4.6)$$

où,  $f(\Omega)$  est la fonction du facteur de forme.

Cette définition est généraliste afin de donner un avantage aux formes spécifiques dans une procédure de segmentation. Elle permet l'utilisation de tous les attributs de forme et de toutes les fonctions de similarité imaginables car il n'y a pas de notion de critère croissant sur la fonction du facteur de forme. Dans un cas spécifique, il reste à définir les images binaires ou les formes sur lesquelles nous calculons la fonction du facteur de forme. Par exemple si nous utilisons la représentation *Max-Tree* pour calculer des ouvertures ultimes par attributs, chaque nœud  $C_h^k$

correspond à une forme  $\Omega$ . Par rapport au facteur de forme nous devons être prudents avec la sélection de mesures qui le compose, pour les raisons suivantes :

- **Temps de calcul** : Comme nous l'avons mentionné, nous utilisons la représentation *Max-Tree* pour calculer l'ouverture ultime par attribut. C'est pourquoi les attributs sont calculés pour chaque nœud. Pour conserver un temps de calcul raisonnable, nous utilisons les attributs de forme les plus simples, parce qu'ils peuvent être facilement estimés au cours de la construction du *Max-Tree*.
- **Invariance à la Translation, à l'Echelle et à la Rotation (ITER)** : Un point très important dans toutes les approches de segmentation basées sur les formes est la robustesse de l'opérateur à la translation, à l'échelle et à la rotation. D'abord, nous analysons l'invariance de l'ouverture ultime  $\theta$ . La robustesse à l'invariance  $\theta$  dépend directement du type d'ouverture  $\gamma_\lambda$  que nous utilisons et de la propriété d'être intrinsèquement multi-échelle. Cela veut dire que si  $\gamma_\lambda$  est complètement invariante (ITER), l'opérateur  $\theta$  le sera de même. Par exemple, l'ouverture ultime par attribut de hauteur est invariante à la translation et à l'échelle ITE. En conséquence, l'invariance du nouvel opérateur  $\theta^\Omega$  dépend de l'invariance de  $\theta$  et de la fonction du facteur de forme  $f(\Omega)$ . Ainsi, l'invariance de  $f(\Omega)$  est fonction de  $\psi_{\forall\kappa}(\Omega)$ . On pourrait dire que l'invariance de  $\theta^\Omega$  est l'intersection des invariants des fonctions dont il dépend. Par exemple si  $\theta$  est ITE et  $\psi_{\forall\kappa}(\Omega)$  est ITER,  $\theta^\Omega$  sera ITE.

## 4.4 Exemple d'une image synthétique

Nous testons cette approche avec l'image synthétique de la Figure 3.6(a) page 42. D'abord nous voulons favoriser les formes rectangulaires. C'est pourquoi nous utilisons la relation de remplissage comme attribut de forme  $\Upsilon_\Omega = \frac{A_\Omega}{Abbox_\Omega}$ , où  $(A_\Omega)$  est la surface de la forme et  $(Abbox_\Omega)$  la surface de la boîte englobante. Cette relation a une dynamique comprise entre 0 et 1, où les formes rectangulaires sans rotation ont une valeur de 1. Ensuite, nous avons imposé des limites de surface afin de valider la forme à favoriser. Nous utilisons une surface limite de 90% de la surface totale de l'image pour pénaliser les plus grandes régions. La fonction du facteur de forme est définie par l'équation 4.7.

$$\begin{aligned}
 f(\Omega) &= 1 + \alpha \psi_{\kappa_1}(\Omega) \psi_{\kappa_2}(\Omega) \\
 \text{où,} \\
 \alpha &= 10, \\
 \psi_{\kappa_1}(\Omega) &= \Upsilon_\Omega \\
 \psi_{\kappa_2}(\Omega) &= \begin{cases} 1 & A_\Omega < 90\% A_I \\ 0 & \text{sinon} \end{cases}
 \end{aligned} \tag{4.7}$$

La Figure 4.1 présente le résultat de la méthode en utilisant l'équation 4.7 comme fonction du facteur de forme. Dans ce cas, le problème du masquage a été résolu et les trois formes ont été détectées. L'exemple illustre l'utilisation d'une limite comme attribut de forme  $\psi_{\kappa_2}(\Omega)$ . L'importance de la limite est remarquable dans cet exemple, parce que le plateau qui masque les structures atteint un facteur de remplissage élevé  $\Upsilon_\Omega \approx 1$ . En utilisant la limite, le facteur de forme est égal à 1, ainsi  $r_\lambda^\Omega = r_\lambda$ . De manière générale l'emploi de limites comme attributs de forme réduit le caractère multi-échelle de  $\theta^\Omega$  aux limites établies. Cette limitation présente des avantages dans certaines applications avec des images réelles.

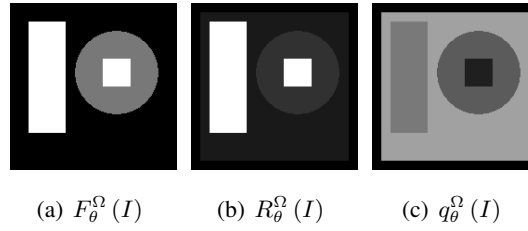


FIGURE 4.1 – Segmentation de l'image synthétique avec  $\theta^\Omega$ . Eq. 4.7 comme fonction du facteur de forme.

Même si  $\Upsilon_\Omega$  favorise principalement les formes rectangulaires sans rotation, le facteur de la structure du cercle est suffisamment élevé pour le démasquer. Si nous ne voulons favoriser que les formes rectangulaires, nous élevons  $\Upsilon_\Omega$  à la puissance trois, comme l'illustre l'équation 4.8. Etant donné que le facteur de remplissage ( $\Upsilon_\Omega$ ) est compris entre 0 et 1,  $(\Upsilon_\Omega)^3$  sera inférieur  $\Upsilon_\Omega$ . De cette manière, le cercle ne sera pas détecté. Le résultat de la segmentation est illustré sur la Figure 4.2(c).

$$\begin{aligned}
 f(\Omega) &= 1 + \alpha \psi_{\kappa_1}(\Omega) \psi_{\kappa_2}(\Omega) \\
 \text{où,} \\
 \alpha &= 10 \\
 \psi_{\kappa_1}(\Omega) &= (\Upsilon_\Omega)^3
 \end{aligned} \tag{4.8}$$

$$\psi_{\kappa_2}(\Omega) = \begin{cases} 1 & A_\Omega < 90\% A_I \\ 0 & \text{sinon} \end{cases}$$

Par ailleurs un autre facteur de forme est mis en œuvre pour favoriser des formes circulaires. Pour la détection des cercles, l'attribut le plus fréquemment utilisé est la circularité. Cet attribut est la relation entre la surface de la forme ( $A_\Omega$ ) et le périmètre ( $L_\Omega$ ) au carré :  $\frac{4\pi A_\Omega}{(L_\Omega)^2}$ . Ainsi, cet attribut est utilisé dans la fonction du facteur de forme avec une dynamique étroite comme l'illustre l'équation 4.9. La Figure 4.3 présente le résultat de  $\theta^\Omega$  en utilisant cette fonction comme facteur.



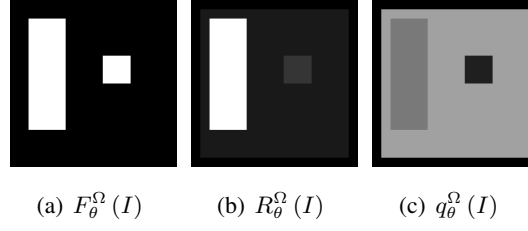


FIGURE 4.2 – Segmentation de l'image synthétique avec  $\theta^{\Omega}$ . Eq. 4.8 comme fonction du facteur de forme.

$$\begin{aligned}
 f(\Omega) &= 1 + \alpha \psi_{\kappa_1}(\Omega) \\
 \text{où,} \\
 \alpha &= 10 \\
 \psi_{\kappa_1}(\Omega) &= \left( \frac{4\pi A_{\Omega}}{(L_{\Omega})^2} \right)^3
 \end{aligned} \tag{4.9}$$

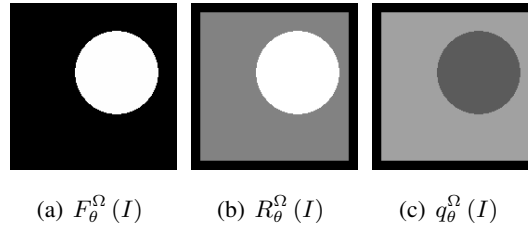


FIGURE 4.3 – Segmentation de l'image synthétique avec  $\theta^{\Omega}$ . Eq. 4.9 comme fonction du facteur de forme.

Comme nous avons pu l'observer dans les trois exemples, nous avons choisi la valeur de  $\alpha = 10$  car le plus petit résidu de la structure à détecter est égal à 25 et le résidu du plateau est égal à 130. Avec ces résidus toutes valeurs de  $\alpha > 5$  peuvent être utilisées en ayant les mêmes résultats de segmentation.

## 4.5 Application : Images de façade

Initialement, nous avons utilisé l'ouverture ultime par attribut de hauteur afin de segmenter les images de façades. Comme image de départ nous avons choisi un gradient couleur HSL [59] inversé. Les résultats sont illustrés sur les Figures 4.4, 4.5 et 4.6 (pages 55, 56 et 57). Dans le premier exemple, l'opérateur  $\theta$  montre un résultat de segmentation acceptable car il a détecté un grand nombre des structures internes. Cependant, quatre des cinq fenêtres de la partie basse de l'immeuble ont été masquées, ainsi que quelques-unes dans les autres étages. Le masquage dans cet exemple est causé par la grande différence de couleur entre la façade et la toiture du bâtiment. Les autres exemples illustrent un problème de masquage plus fort par rapport à l'exemple précédent, parce que la plupart des

structures de la façade n'ont pas été détectées. Le problème se produit car le contraste entre le ciel et le bâtiment est plus grand que le contraste entre le mur et les fenêtres. La plupart des images urbaines contiennent de l'information du ciel et/ou des différences de couleur entre le toit et le mur ; c'est pourquoi l'utilisation de l'ouverture ultime sur ce type d'images peut être fortement affectée par le problème de masquage dû aux structures imbriquées.

Dans le but de segmenter les fenêtres dans les images de façade, nous analysons leurs caractéristiques particulières. Mayer et Reznik [103] décrivent la plupart des fenêtres comme une forme particulièrement rectangulaire et dont la relation hauteur ( $h$ ) largeur ( $w$ ), relation d'aspect, est généralement comprise entre 0.2 et 5.0. En utilisant ces caractéristiques nous pouvons générer des fonctions de similarité de fenêtres et ainsi la fonction du facteur de forme  $f(\Omega)$ .

Afin d'exploiter la caractéristique rectangulaire des fenêtres, nous utilisons le facteur de remplissage comme un attribut de forme  $\kappa_1 = \Upsilon$ . Cet attribut est élevé au carré ( $\varsigma = 2$ ) pour faire une fonction de similarité de dynamique étroite et ainsi pénaliser les formes non-rectangulaires ou les rectangles avec des trous. Le deuxième attribut de forme utilisé est la relation d'aspect ( $h/w$ ). Cet attribut est trop sensible aux petites variations de la largeur et de plus, si nous l'utilisons directement comme fonction de similarité, elle ne vérifie pas les conditions présentées dans la Définition 10. C'est pour cela que nous avons défini la relation d'aspect de la manière suivante :  $\aleph = \frac{\min(h,w)}{\max(h,w)}$ . Ainsi, nous tronquons  $\aleph$  supérieure à 0.2 et nous assignons une dynamique large  $\varsigma = 0.5$  pour créer la fonction de similarité. Nous avons également ajouté des contraintes par rapport à la hauteur et à la largeur de l'image,  $H_I$  et  $W_I$  respectivement ; étant donné que nous travaillons sur des images comprenant une façade entière, la hauteur d'une fenêtre est rarement supérieure à un tiers de  $H_I$  et de même pour la largeur. Finalement, il nous reste à définir la valeur  $\alpha$ . Cette valeur a été déterminée de manière empirique car après une analyse de l'image de gradient, on a constaté que le contraste des structures internes est 10 fois inférieur au gradient le plus fort de l'image. C'est pourquoi nous avons choisi  $\alpha = 9$ , i.e.  $1 + \alpha = 10$ . L'équation 4.10 illustre la fonction du facteur de forme finale pour la détection de fenêtres.

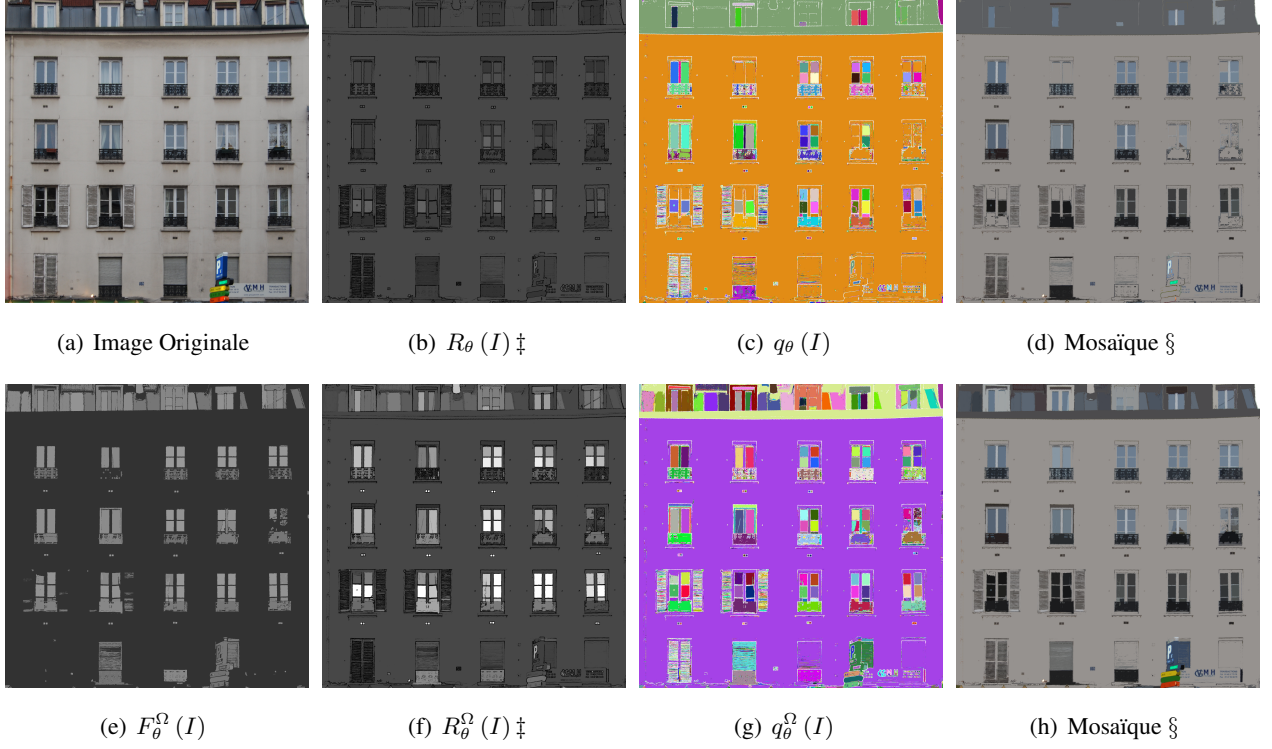


FIGURE 4.4 – (a) Image Originale, (b)-(c)  $R_\theta(I)$  et  $q_\theta(I)$  issues de l'ouverture ultime, (d) Image Mosaïque sur  $q_\theta(I)$ . (e)-(g)  $F_\theta^\Omega(I)$ ,  $R_\theta^\Omega(I)$  et  $q_\theta^\Omega(I)$  issues de l'ouverture ultime avec l'information de forme, (h) Image Mosaïque sur  $q_\theta^\Omega(I)$ .  $\ddagger$  Correction gamma de 2. L'image mosaïque (§) est la valeur moyenne de l'image originale sur la partition générée par la ligne de partage des eaux du gradient en utilisant comme marqueur  $q_\theta$ .

$$\begin{aligned}
 f(\Omega) &= 1 + \alpha \psi_{\kappa_1}(\Omega) \psi_{\kappa_2}(\Omega) \psi_{\kappa_3}(\Omega) \\
 \text{où,} \\
 \alpha &= 9 \\
 \psi_{\kappa_1}(\Omega) &= (\Upsilon_\Omega)^2 \\
 \psi_{\kappa_2}(\Omega) &= \begin{cases} (\aleph_\Omega)^{0.5} & \aleph_\Omega > 0.2 \\ 0 & \text{sinon} \end{cases} \\
 \psi_{\kappa_3}(\Omega) &= \begin{cases} 1 & h_\Omega < \frac{H_I}{3} \wedge w_\Omega < \frac{W_I}{3} \\ 0 & \text{sinon} \end{cases}
 \end{aligned} \tag{4.10}$$

Dans les Figures 4.4, 4.5 et 4.6 (pages 55, 56 et 57), nous pouvons observer les résultats des opérateurs  $\theta$  et  $\theta^\Omega$ . Nous présentons aussi une image mosaïque dans chacun des exemples pour remarquer les différences entre les

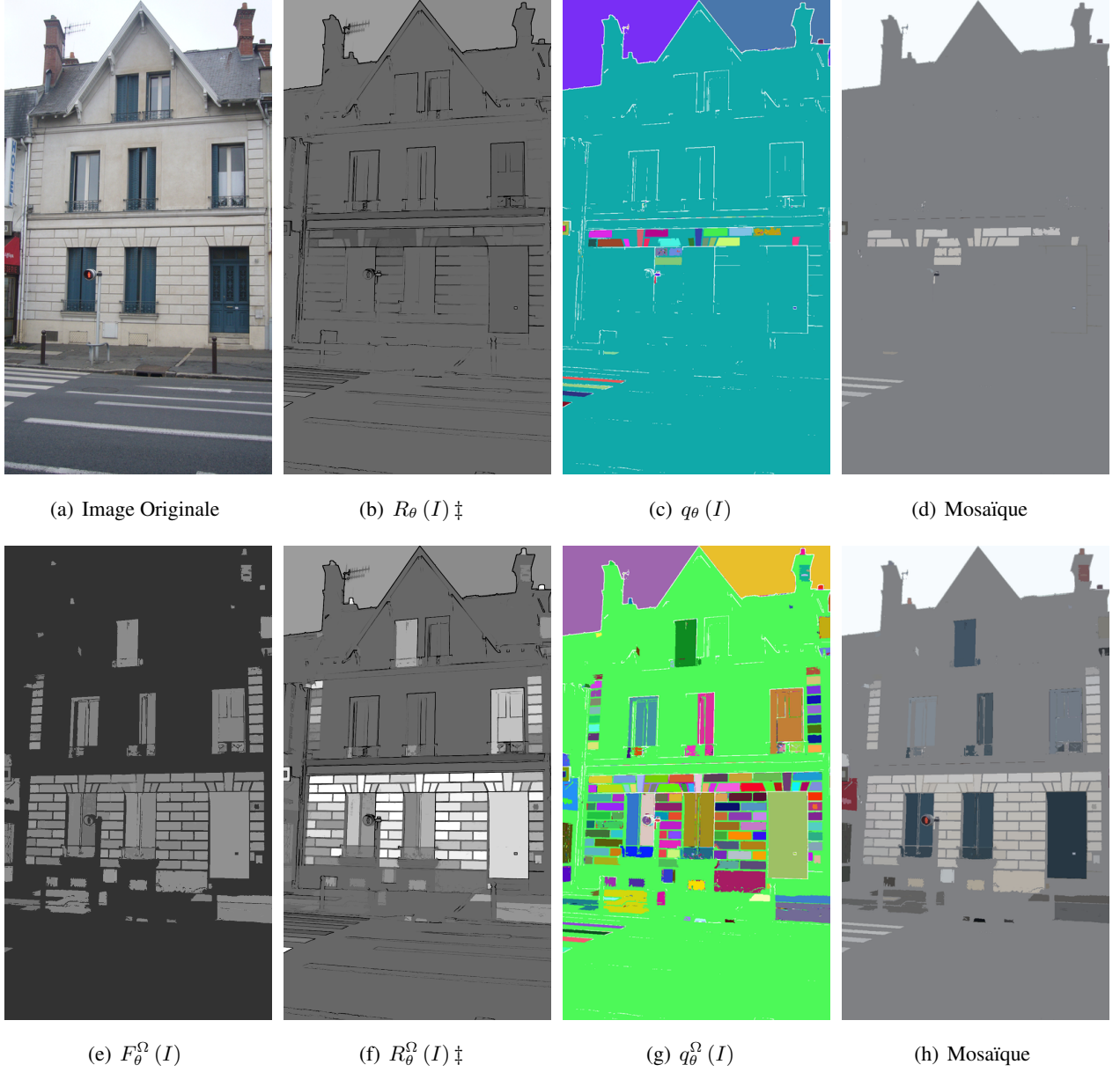


FIGURE 4.5 – (a) Image Originale, (b)-(c)  $R_\theta(I)$  et  $q_\theta(I)$  issues de l'ouverture ultime, (d) Image Mosaïque sur  $q_\theta(I)$ . (e)-(g)  $F_\theta^\Omega(I)$ ,  $R_\theta^\Omega(I)$  et  $q_\theta^\Omega(I)$  issues de l'ouverture ultime avec l'information de forme, (h) Image Mosaïque sur  $q_\theta^\Omega(I)$ .  $\ddagger$  Correction gamma de 2.

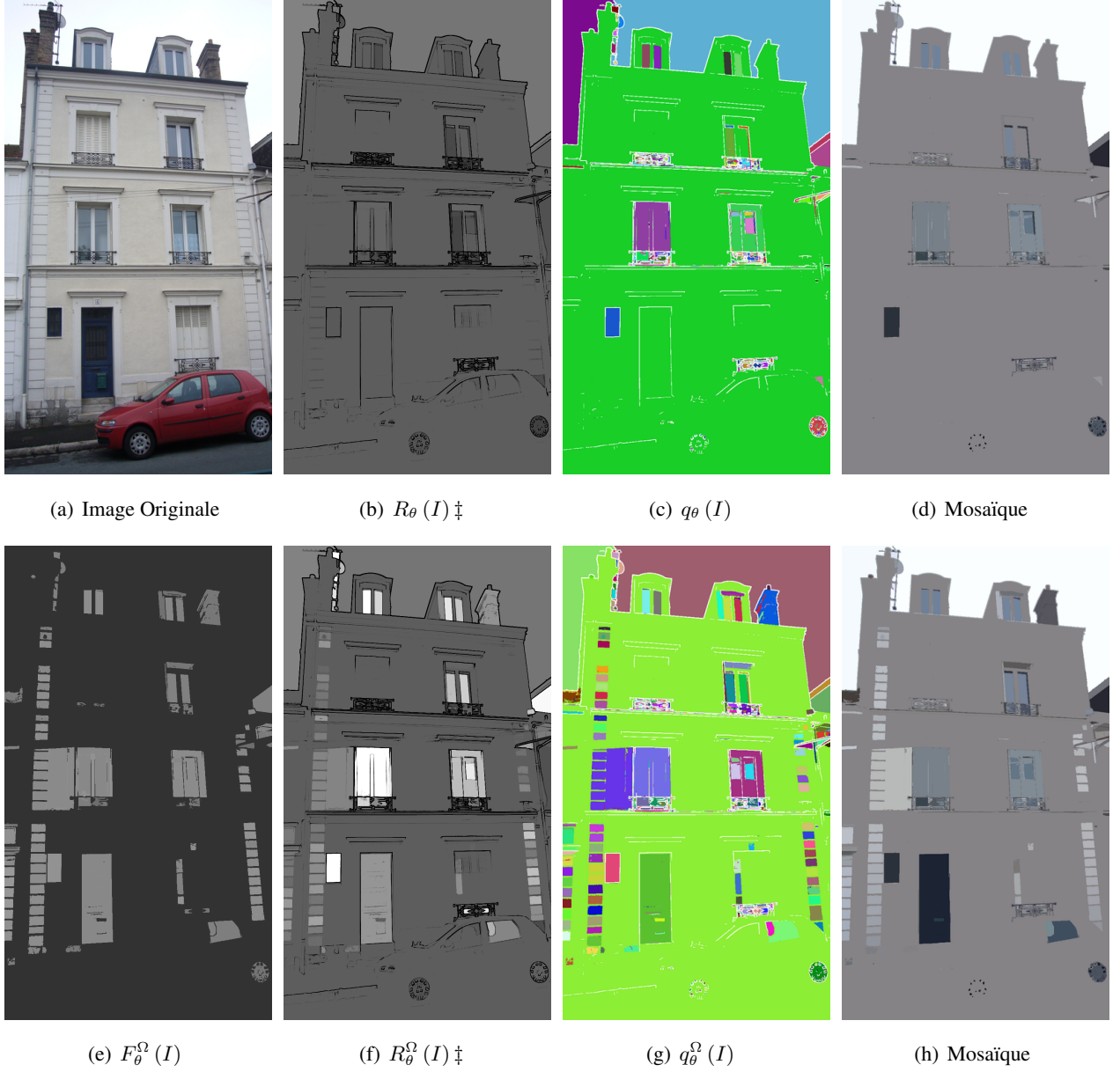


FIGURE 4.6 – (a) Image Originale, (b)-(c)  $R_\theta(I)$  et  $q_\theta(I)$  issues de l'ouverture ultime, (d) Image Mosaïque sur  $q_\theta(I)$ . (e)-(g)  $F_\theta^\Omega(I)$ ,  $R_\theta^\Omega(I)$  et  $q_\theta^\Omega(I)$  issues de l'ouverture ultime avec l'information de forme, (h) Image Mosaïque sur  $q_\theta^\Omega(I)$ .  $\ddagger$  Correction gamma de 2.

deux méthodes. Comme nous pouvons l'observer, notre approche présente de meilleurs résultats dans les trois cas. Même si la fonction du facteur a été conçue pour favoriser la détection de fenêtres, toutes les structures qui ont les mêmes caractéristiques seront aussi artificiellement affectées par la fonction de forme. Ce comportement est bien perçu sur des applications des images de façade car d'autres éléments comme les portes, les briques et les stores, ont des caractéristiques de forme similaires aux fenêtres et seront donc favorisés. Mise à part les briques, les autres structures détectées sont des éléments utiles pour la modélisation d'une façade, c'est pourquoi ce comportement n'est pas considéré comme problématique pour l'application.

Les résultats sont exposés de la manière suivante :

- Dans le premier exemple dix fenêtres ont été récupérées et seulement deux autres ont été manquées dans la détection et même des détails du toit sont ressortis.
- Dans le deuxième exemple on a détecté les fenêtres et la texture de la façade car elle est faite en briques.
- Dans le troisième exemple, la porte a été segmentée ainsi que les détails de la façade. Pour cette image résultat, deux fenêtres avec store n'ont pas été détectées (2ème étage à gauche et rez-de-chaussée à droite).

Pour mieux comprendre le problème des fenêtres manquantes, voyons les images de la Figure 4.7, page 59. Nous avons déjà enregistré sur une image  $F_{\theta}^{\Omega}(I)$  la valeur de la fonction de forme  $f(\Omega)$  quand le résidu maximal ( $R_{\theta}(I)$ ) est généré. Simultanément, nous enregistrons la valeur maximale de  $f(\Omega)$  sur l'image  $\widehat{F}_{\theta}^{\Omega}(I)$ . Sur cette image on peut estimer la valeur du facteur dans les régions de l'image. Dans l'exemple de la Figure 4.6(g) deux fenêtres n'ont pas été détectées. En observant la Figure 4.7(d), nous pouvons apprécier que les deux fenêtres ont un fort facteur de forme ; cependant en observant la Figure 4.7(c), nous constatons le faible contraste entre les fenêtres et le mur, causé par la couleur du store, similaire à celle du mur. Une solution à ce problème est l'augmentation du paramètre  $\alpha$ , néanmoins ce changement produira aussi plus de régions de bruit.

## 4.6 Discussion

Dans ce chapitre nous avons présenté une première stratégie pour réduire les problèmes de l'ouverture ultime. Le nouvel opérateur  $\theta^{\Omega}$  associe l'information de forme *a priori*, dans une application donnée, en utilisant une fonction de similarité. La méthode introduit l'information de forme comme un facteur multiplicatif dans la définition de l'ouverture ultime. L'opérateur présente deux images de sortie additionnelles par rapport à la version classique de l'ouverture ultime :  $F_{\theta}^{\Omega}(I)$  et  $\widehat{F}_{\theta}^{\Omega}(I)$ .  $F_{\theta}^{\Omega}(I)$  enregistre la valeur de la fonction du facteur de forme  $f(\Omega)$  quand le résidu maximal ( $R_{\theta}(I)$ ) est généré et  $\widehat{F}_{\theta}^{\Omega}(I)$  la valeur maximale de  $f(\Omega)$ . Cette nouvelle définition d'opérateur  $\theta^{\Omega}$  permet l'utilisation de tous les attributs imaginables car la fonction  $f(\Omega)$  reste indépendante de l'ouverture ultime même. Néanmoins, dans ce chapitre nous utilisons seulement des attributs simples afin de garder un temps de calcul raisonnable nous appuyant sur la représentation *Max-Tree* qui est employée pour le calcul de l'ouverture ultime.

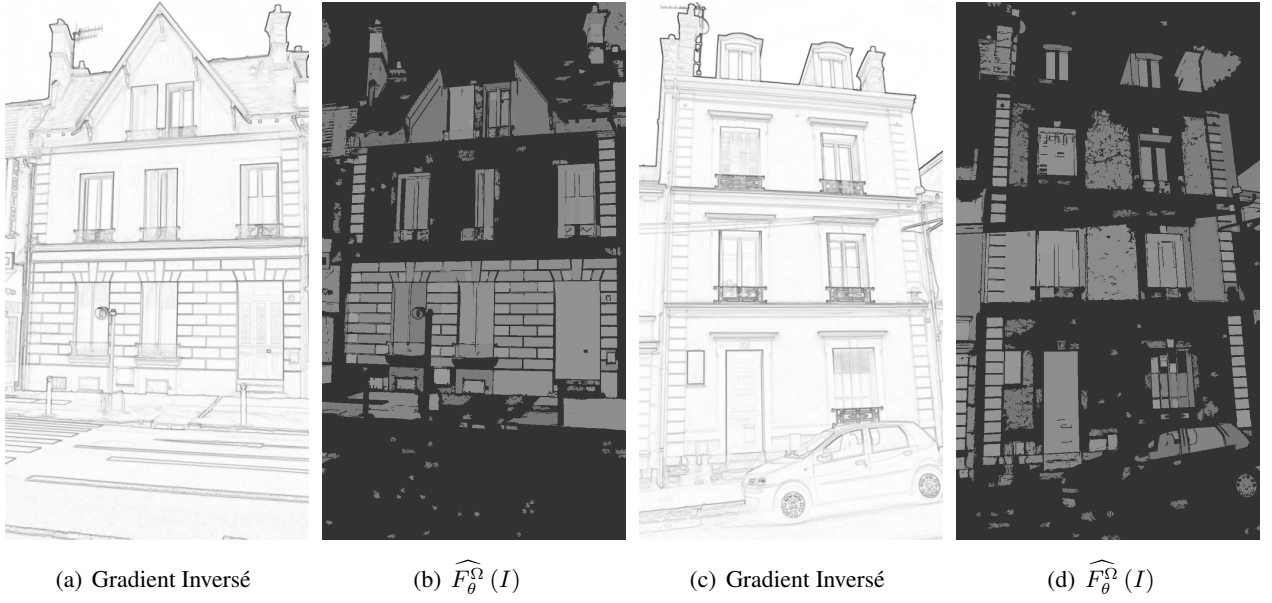


FIGURE 4.7 – (a-b) Images de l'exemple de la Figure 4.5. (c-d) Images de l'exemple de la Figure 4.6.

Par rapport à la fonction de similarité, certaines conditions doivent être vérifiées, telles que la réponse normalisée et non négative, pour l'utilisation dans notre approche. Toutefois, la sortie d'un processus d'apprentissage, en considérant ces conditions, pourrait être également introduite comme fonction du facteur. Par ailleurs, même si nous avons sacrifié le côté non paramétrique de l'ouverture ultime, l'approche reste facilement paramétrable pour son utilisation. Par exemple, une application avec des images de façade a été illustrée, en montrant la puissance de l'approche. L'information de forme utilisée comme *a priori* de la méthode a été facilement sélectionnée. Les résultats de segmentation montrent la diminution du problème de masquage au niveau des fenêtres et des éléments de caractéristiques similaires. Des résultats d'autres applications sur des images réelles seront illustrés au Chapitre 7.

# Ouverture Ultime Hiérarchique

## Sommaire

<b>5.1</b>	<b>Introduction</b>	<b>60</b>
<b>5.2</b>	<b>Hiérarchisation de l'ouverture ultime</b>	<b>61</b>
5.2.1	Niveau optimal	66
5.2.2	Mise en œuvre sur le <i>Max-Tree</i>	66
5.2.3	Critère hiérarchique	67
<b>5.3</b>	<b>Application : Images de façade</b>	<b>68</b>
<b>5.4</b>	<b>Discussion</b>	<b>70</b>

## 5.1 Introduction

De nombreuses approches hiérarchiques ont été utilisées pour segmenter des images [13, 58, 76, 108, 117, 137]. Nous pouvons identifier deux catégories principales : des approches ascendantes (*bottom-up*) et descendantes (*top-down*). Les approches ascendantes, à partir d'une sur-segmentation, fusionnent des régions avec un critère donné. Par exemple, la segmentation hiérarchique basée sur la ligne de partage des eaux (*LPE*), présentée par Meyer dans [108], est une méthode permettant d'obtenir une hiérarchie de partitions de l'image traitée en s'appuyant sur la structure d'un graphe associé. Les approches descendantes commencent le processus à partir d'une segmentation grossière (sous-segmentation) et puis les régions sont raffinées au fur et à mesure. Le raffinement peut être total, toutes les régions sont re-segmentées, ou local, en utilisant un critère afin de choisir certaines régions. Une approche descendante a été présentée par Salembier dans [137]. La méthode est basée sur l'extraction de marqueurs selon des critères de contraste et de taille. La localisation de ces marqueurs détermine les régions qui devront être subdivisées. Une LPE, contrainte à partir de marqueurs sélectionnés, produit la partition de niveau hiérarchique suivant.

Tel que l'illustre le Chapitre 3, l'ouverture ultime présente des problèmes qui peuvent être vus comme un problème de sous-segmentation. C'est pourquoi, nous proposons une approche descendante avec la hiérarchisation de l'ouverture ultime. Tout d'abord dans la section 5.2, nous introduisons deux notions importantes : l'opérateur de mas-



quage et l'ouverture ultime géodésique. Puis, nous définissons l'ouverture ultime hiérarchique de niveau  $N$ , en utilisant un critère  $T$ . La section 5.3 illustre une application sur des images de façade. D'autres exemples d'utilisation de cet opérateur seront présentés au Chapitre 7. Pour finir, nous discuterons les avantages et désavantages de notre approche dans la dernière section.

## 5.2 Hiérarchisation de l'ouverture ultime

La hiérarchisation d'un opérateur résiduel consiste à déterminer les résidus d'un niveau donné ou les  $N$ -résidus maximaux comme illustré dans [88]. Pour analyser la possible hiérarchisation de l'opérateur, nous avons identifié plusieurs questions à étudier : parmi les  $N$  niveaux de résidus, quel est le niveau optimal  $N$  ? Quelle est la sortie finale de l'opérateur combinant les objets qui apparaissent à différents niveaux ? Est-ce que le niveau optimal  $N$  peut être associé à un critère hiérarchique  $T$  ? Si oui, quel est le critère à utiliser ?

Ces questions seront abordées afin de définir l'ouverture ultime hiérarchique et pour cela, nous commencerons par introduire l'ouverture ultime géodésique d'une image  $I$ , en nous appuyant sur un opérateur de masquage. Tout d'abord nous allons définir la notion de masquage d'une image  $I$  par une image binaire  $X$ .

**Définition 12.** Soit  $I$  une image à niveaux de gris et  $X$  une image binaire, l'opérateur de masquage  $\mathcal{M}_X(I)$  est décrit par :

$$\mathcal{M}_X(I(x)) = \begin{cases} I(x) & X(x) = \text{vrai} \\ 0 & \text{sinon} \end{cases} \quad (5.1)$$

Rappelons la définition de l'opérateur classique  $\theta$ , où  $q_\theta(I)$  est l'indicatrice (information de taille) et  $R_\theta(I)$  est la transformée (information de contraste). Pour alléger la notation, nous dénotons  $q_\theta(I) = q_\theta$  et  $R_\theta(I) = R_\theta$ .

$$\theta(I) : I \xrightarrow{\theta} (R_\theta, q_\theta) \quad (5.2)$$

**Définition 13.** Nous définissons l'ouverture ultime géodésique de  $I$ ,  $\theta_X(I)$ , pour un masque binaire  $X$ , comme l'ouverture ultime de  $I$  appliquée sur le domaine de  $X$ . L'opérateur  $\theta_X(I)$  est donc défini à l'intérieur des composantes connexes de  $X$  que nous désignerons désormais par l'espace géodésique<sup>a</sup>  $X$ .

<sup>a</sup>. Ce terme a aussi été utilisé par Beucher dans sa thèse [12] pour la définition de la distance géodésique.

Le masque  $X$  sert donc comme domaine de support de l'ouverture ultime et il contraint la connexité entre pixels. Le lecteur doit faire attention à la mise en œuvre de cet opérateur car il ne s'agit pas de masquer l'image d'entrée pour appliquer l'opérateur  $\theta(\mathcal{M}_X(I))$ , ni de masquer la sortie de l'opérateur  $\mathcal{M}_X(\theta(I))$ . Afin d'éclaircir ces

différences, un exemple de l'ouverture ultime géodésique par attribut de hauteur est illustré dans la Figure 5.1. Ainsi, la structure la plus contrastée de  $I$  dans l'espace géodésique  $X$  est le carré, qui a été détecté par  $\theta_X(I)$  (voir Figure 5.1(c)). Dans le cas de la Figure 5.1(d), le fait de masquer l'image d'entrée introduit des structures contrastées dans l'image (liées au masque lui-même) qui n'existaient pas sur l'image de départ, en produisant une segmentation erronée. Finalement, le fait de masquer l'image de sortie n'évite pas que des structures fortement contrastées, en dehors de l'espace géodésique, éliminent des structures d'intérêt à l'intérieur de l'espace géodésique. La Figure 5.1(e) montre que toutes les structures ont été éliminées par le résidu du grand plateau<sup>1</sup>.

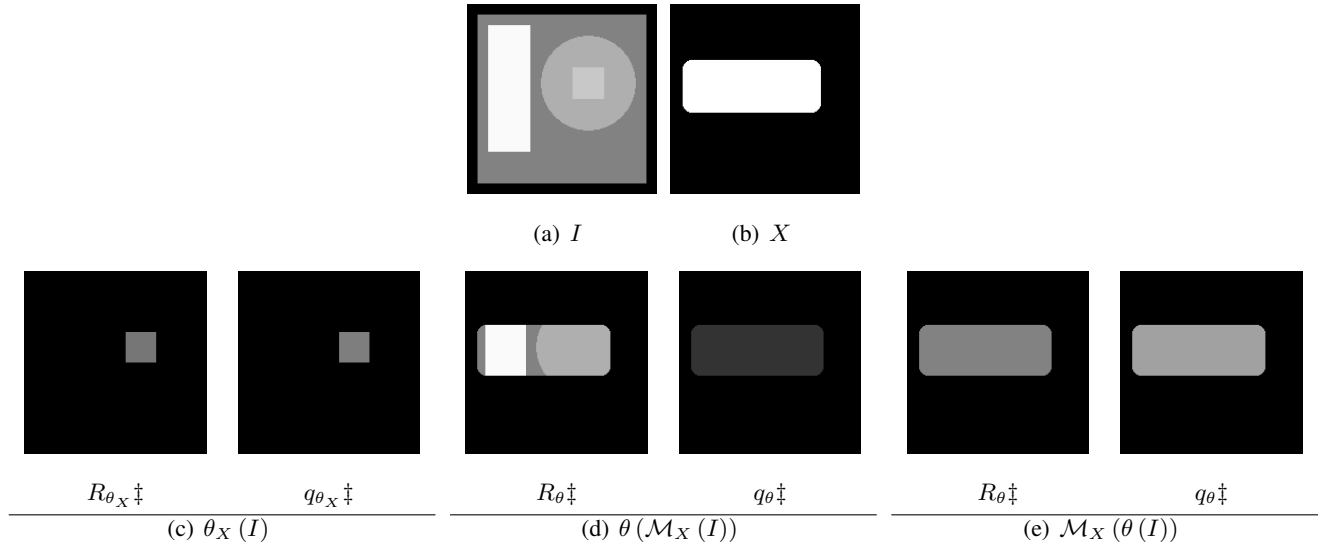


FIGURE 5.1 – (a) Image  $I$ , (b) Masque  $X$ . (c) Ouverture ultime géodésique de  $I$  en  $X$ . (d) Ouverture ultime du masquage de  $I$  en  $X$ . (e) Ouverture ultime de  $I$  masquée par  $X$ .  $\ddagger$  Correction gamma de 3

Nous allons étendre la définition de  $\theta_X(I)$  pour utiliser une image d'étiquettes  $L$  comme espace géodésique.

**Définition 14.** Soit  $I$  une image à niveaux de gris et  $L$  une image d'étiquettes, une ouverture ultime géodésique  $\theta_L(I)$  est décrite par :

$$\theta_L(I) = \bigvee_{i=1}^l \theta_{L_i}(I) \quad (5.3)$$

où  $l$  est le nombre d'étiquettes,  $L_i$  est l'image binaire,  $L_i(x) = \{x \in E | L(x) = i\}$  et  $\theta_{L_i}(I)$  est l'ouverture ultime géodésique pour un masque binaire  $L_i$  (voir définition 13.).

La notion de hiérarchie de l'ouverture ultime est basée sur l'idée de re-segmenter, dans une approche *top-down*, en utilisant l'opérateur de manière géodésique sur certaines régions de la segmentation précédente. Nous analyserons

1. Le lecteur peut se référer à la page 42 et la Figure 3.6 pour observer le problème dû aux structures imbriquées sur cette image synthétique.

l'image  $q_\theta$  comme sortie de segmentation de l'opérateur, étant donné que cette image correspond à la taille des ouvertures qui ont produit les problèmes de masquage ; cependant d'autres images, comme  $R_\theta$ , peuvent être aussi utilisées. La sélection de régions à re-segmenter se fait à l'aide d'un critère<sup>2</sup>  $T(C) \rightarrow \{vrai, faux\}$  comme critère hiérarchique sur les zones plates de  $q_\theta$  différentes de zéro.

**Définition 15.** Soit  $I$  une image à niveaux de gris et  $L$  l'image d'étiquettes de zones plates de  $q_\theta$  différentes de zéro. L'image d'étiquettes  $V$  est définie comme l'ensemble de régions de  $L$  qui vérifient le critère  $T$  :

$$V(x) = \bigvee_{i=1}^l \begin{cases} i & T(L_i(x)) = vrai \\ 0 & sinon \end{cases} \quad (5.4)$$

Dans une procédure hiérarchique de l'ouverture ultime avec un critère  $T$ , nous désignons l'opérateur classique  $\theta$  comme niveau zéro de la hiérarchie :

$$\theta_T^0(I) = \theta(I) : I \xrightarrow{\theta} (R_\theta^0, q_\theta^0) \quad (5.5)$$

Pour le niveau zéro, le critère  $T$  n'est pas utilisé. Ainsi, nous utilisons le critère à partir du premier niveau de hiérarchie sur l'image d'étiquettes  $L^0$  des zones plates de  $q_\theta^0$  différentes de zéro. Nous étudions le critère  $T(L^0) \rightarrow V^0$ , afin de déterminer les régions sur lesquelles nous appliquons l'opérateur géodésique. Ensuite, nous appliquons l'opérateur géodésique  $\theta_{V^0}$  en utilisant  $V^0$  comme masque géodésique (voir Définition 14).

$$\theta_{V^0}(I) : I \xrightarrow{\theta_{V^0}} (R_{\theta_{V^0}}, q_{\theta_{V^0}}) \quad (5.6)$$

Dans cette étape, nous avons appliqué l'opérateur sur certaines régions de l'image. Néanmoins, la sortie finale du premier niveau de hiérarchie combine l'ensemble de régions du niveau zéro qui ne vérifient pas le critère  $\bar{V}^0$ , et la sortie de l'opérateur géodésique sur l'espace géodésique  $V^0$ , comme l'illustre l'équation 5.7.

$$\theta_T^1(I) = \mathcal{M}_{\bar{V}^0}(\theta_T^0(I)) \vee \theta_{V^0}(I) \quad (5.7)$$

Nous réitérons la procédure pour déterminer le deuxième niveau de hiérarchie, avec  $T(L^1) \rightarrow V^1$ , donnant comme sortie finale de l'opérateur hiérarchique :

$$\theta_T^2(I) = \mathcal{M}_{\bar{V}^0}(\theta_T^0(I)) \vee \mathcal{M}_{\bar{V}^1}(\theta_T^1(I)) \vee \theta_{V^1}(I) \quad (5.8)$$

Finalement, la procédure peut être itérée pour tout  $k$ -ième niveau jusqu'au niveau de hiérarchie  $N$  donné. Ainsi, l'ouverture ultime hiérarchique  $\theta_T^N(I)$  est présentée dans la Définition 16.

2. Voir la Définition 7 page 35 pour un critère.

**Définition 16.** Soit  $I$  une image à niveaux de gris,  $T$  un critère hiérarchique et  $N$  le niveau de la hiérarchie, une ouverture ultime hiérarchique  $\theta_T^N(I)$  est décrite par :

$$\theta_T^N(I) = \left( \bigvee_{k=0}^{N-1} \mathcal{M}_{\overline{V}^k} \left( \theta_T^k(I) \right) \right) \vee \theta_{V^{N-1}}(I) \quad (5.9)$$

où

$\theta_T^k(I) : I \xrightarrow{\theta_T^k} (R_{\theta_T^k}, q_{\theta_T^k})$  est l'ouverture ultime hiérarchique du  $k$ -ième niveau,

$\theta_{V^k}(I) : I \xrightarrow{\theta_{V^k}} (R_{\theta_{V^k}}, q_{\theta_{V^k}})$  est l'ouverture ultime géodésique de  $I$  sur l'image d'étiquettes  $V^k$ ,

$T(L^k) \rightarrow V^k$  est l'image des étiquettes qui vérifient le critère  $T$ , au  $k$ -ième niveau,

$L^k$  sont les zones plates de  $q_{\theta_T^k}$  différentes de zéro, et

$\mathcal{M}_{\overline{V}^k}()$  est l'opérateur de masquage sur les régions qui ne vérifient pas le critère (voir Définition 5.1).

La Figure 5.2 présente un exemple d'ouverture ultime hiérarchique. L'ouverture utilisée est une ouverture par attribut de hauteur. Nous avons utilisé comme critère hiérarchique la région la plus grande de l'indicatrice du niveau précédent. Dans l'exemple nous avons analysé trois niveaux de hiérarchie. L'image  $I$  sur laquelle nous appliquons l'opérateur est l'inverse du gradient couleur de la Figure 5.2(a). En comparant les Figures 5.2(b) et 5.2(c), nous pouvons observer que le résultat hiérarchique présente une meilleure segmentation car il a plus de détails au niveau de la façade. Le déroulement de la hiérarchie est décrit de la manière suivante :

- Niveau 0 : l'opérateur a segmenté les deux bâtiments et la moitié de la rue en une seule région.
- Niveau 1 : à partir de la segmentation de sortie nous vérifions le critère choisi. Ainsi, la région la plus grande où les bâtiments sont fusionnés, est re-segmentée. La segmentation dans ce premier niveau présente les bâtiments séparés, mais l'un des deux reste collé à la rue.
- Niveau 2 : la région qui reste fusionnée est le nouveau domaine d'analyse. La sortie de l'opérateur, dans le niveau hiérarchique deux, sépare le bâtiment de la rue et il détecte aussi les étages.
- Niveau 3 : le nouveau domaine d'analyse est le deuxième étage du bâtiment. Sa segmentation permet la séparation de la façade latérale et frontale du bâtiment.

Avec ce critère et un nombre infini de niveaux, la sortie finale de l'opérateur sera complètement nulle, parce qu'il y aura toujours une région à subdiviser qui vérifie le critère, ainsi l'image d'étiquettes qui vérifie le critère  $V^k$  sera nulle seulement quand l'indicatrice sera nulle. Quand l'ouverture ultime est appliquée à l'intérieur d'une zone plate aucun résidu d'ouverture n'est possible et l'indicatrice sera nulle.



FIGURE 5.2 – (a) Image de façade, (b) indicatrice  $q_\theta$  de l'opérateur classique  $\theta(I)$  et niveau zéro d'hiérarchie, (c)  $q_{\theta_T^3}$  indicatrice de l'opérateur hiérarchique  $\theta_T^3(I)$ , (d)-(f) images d'étiquettes  $V^k$  qui vérifient le critère et (g)-(i) indicatrice  $q_{\theta_{V^k}}$  de l'opérateur géodésique  $\theta_{V^k}(I)$ .

### 5.2.1 Niveau optimal

Dans le cas d'une hiérarchie classique, le critère  $T$  n'est pas utilisé pour déterminer les régions à diviser dans le niveau suivant,  $T(C) \rightarrow \{vrai\}$  donc  $V^k = L^k$ , et toutes les régions sont re-segmentées. En conséquence, toutes les régions de l'image sont affectées par le même  $k$ -ième niveau. Ainsi, en appliquant plusieurs fois une hiérarchie dans un opérateur résiduel, la sortie devient zéro. Pour l'ouverture ultime hiérarchique  $\theta^{N=\infty}(I) = \emptyset$ . Par ailleurs, dans certains cas, les structures à détecter peuvent se trouver dans différents niveaux de hiérarchie. Par exemple, la route peut apparaître au premier niveau de hiérarchie et les fenêtres au deuxième. Si on impose le même niveau de hiérarchie pour l'ensemble de l'image soit les fenêtres manquent (si on choisit le premier niveau), soit la route est sur-segmentée (si on choisit le deuxième niveau). C'est pourquoi nous proposons une re-segmentation selon un critère. En utilisant un critère, nous combinons des régions de différents niveaux hiérarchiques dans le résultat final. Donc, si nous utilisons une valeur de  $N$ , cette valeur sera le niveau maximum qu'il est possible d'atteindre. Ainsi, la valeur  $N$  n'est pas forcément nécessaire si  $V^\infty \rightarrow \emptyset$  ce qui n'implique pas forcément  $L^\infty \rightarrow \emptyset$ . De cette manière, nous pourrions descendre de niveau de hiérarchie jusqu'à l'idempotence de l'opérateur,  $\theta_T^k(I) = \theta_T^{k-1}(I)$ . D'ailleurs ces niveaux de hiérarchie seront les niveaux optimaux par région. Par conséquent, nous avons éliminé la contrainte de la sélection du nombre de niveaux car ils sont fonction du critère  $T$ . De plus, le choix d'un critère reste plus intuitif et simple que la sélection d'un niveau, pour une application donnée. Cependant, ce choix restera un paramètre important pour la réussite du nouvel opérateur et pourra dépendre de l'application. Ce critère  $T$  peut être associé à une fonction d'énergie de la même façon que L. Guigues l'a présenté dans sa thèse [58]. L'idée est de chercher une partition  $P$  de l'image qui minimise une certaine énergie totale  $E(P)$  de la forme :

$$E(P) = \sum_{R \in P} E_D(R) + kE_C(R) \quad (5.10)$$

Où le terme  $k$  pondère la contribution des énergies,  $E_D$  mesure la distance entre la partition et l'image, et  $E_C$  pénalise les segmentations trop fines. Par ailleurs, nous utiliserons une fonction analogue à cette énergie dans la section 7.2.1, afin d'évaluer les différentes segmentations proposées dans cette thèse.

### 5.2.2 Mise en œuvre sur le *Max-Tree*

Le *Max-Tree* est une représentation de l'image intrinsèquement hiérarchique. En effet, la structure racine - feuilles permet la mise en œuvre facile de l'opérateur  $\theta_T(I)$  basé sur un critère  $T$  car l'évaluation du critère doit être réalisée nœud par nœud. La Figure 5.3 illustre la représentation *Max-Tree* de l'image de la Figure 5.3(a). La Figure illustre également un tableau des attributs associés à chaque nœud du *Max-Tree*. Le résultat d'une ouverture ultime par hauteur  $\theta$  est illustré sur la Figure 3.4, page 39.

Dans la définition de l'opérateur  $\theta_T$ , nous descendons d'un niveau de hiérarchie si une région vérifie le critère  $T$ . Sur le *Max-Tree*, nous ne prenons pas en compte les nœuds qui vérifient le critère et nous assignons  $r(C_h^k) = 0$ ,

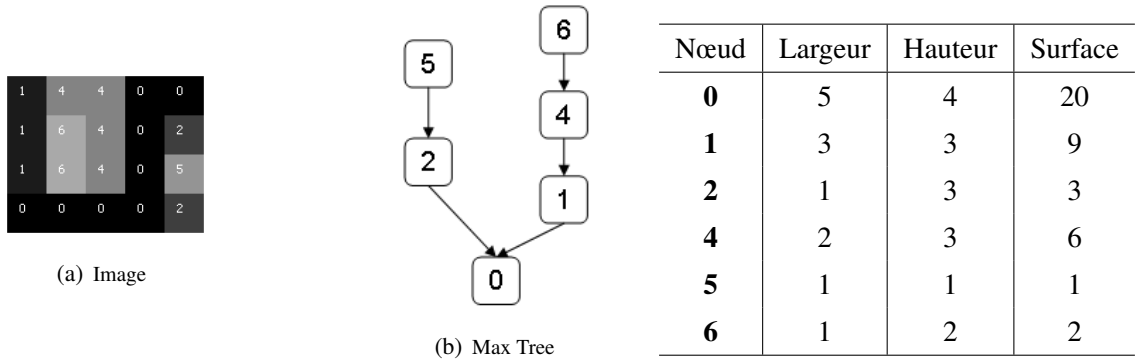
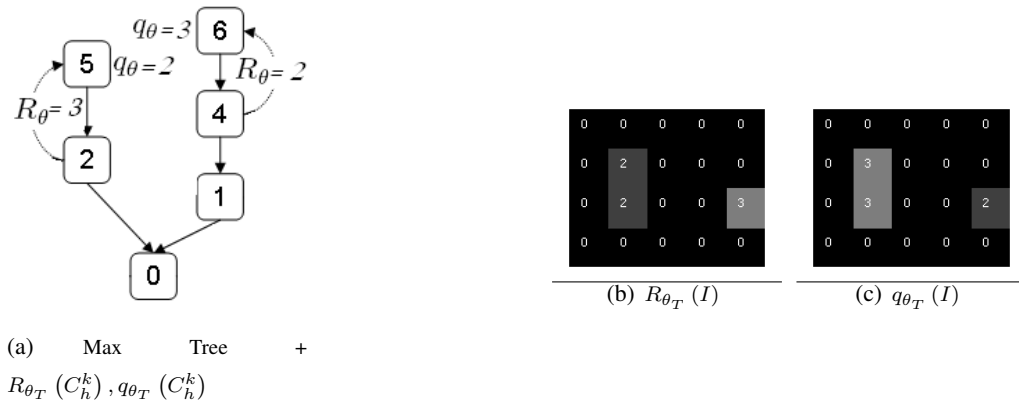


FIGURE 5.3 – (a) Image Synthétique, (b) Max-Tree et tableau d'attributs

c'est-à-dire que le calcul des résidus se fait seulement pour les nœuds qui ne vérifient pas le critère hiérarchique. Ainsi, le calcul de  $R_\theta$  et  $q_\theta$  est complètement identique que celui de la version classique.

Pour l'exemple de la figure synthétique, en utilisant comme critère hiérarchique  $T : \text{Surface}(C_h^k) > 2$ , nous obtenons le résultat de la Figure 5.4. Seuls les nœuds 5 et 6 ne vérifient pas le critère et ils sont donc les seuls nœuds sur lesquels nous calculons les résidus. Ces résidus sont les résidus maximaux et ils font partie de l'image sortie  $R_\theta$ . Le calcul de la hiérarchie est trivial dans cet exemple. Afin d'illustrer la puissance de la méthode, un exemple sur une application d'images de façade sera illustré dans la section 5.3.


 FIGURE 5.4 – (a)  $R_{\theta_T}(C_h^k), q_{\theta_T}(C_h^k)$  et les images résultats de : (b)  $R_{\theta_T}(I)$ , et (c)  $q_{\theta_T}(I)$  de l'ouverture ultime hiérarchique par hauteur.  $T : \text{Surface}(C_h^k) > 2$ 

### 5.2.3 Critère hiérarchique

L'utilisation de la taille d'arrêt, comme la dernière ouverture utilisée, est un cas particulier de l'ouverture ultime hiérarchique par attribut, lorsque le critère choisi dans la hiérarchie dépend seulement du même attribut que

l'ouverture. De plus, nous pouvons utiliser tous les critères qui dépendent d'attributs géométriques comme ceux présentés dans la Figure 3.1, page 36, ou des fonctions de forme comme nous l'avons décrit dans la section 4.2. Cependant, afin de garder un temps de calcul raisonnable, nous utiliserons des critères qui dépendent d'attributs simples tels que la largeur, la hauteur et la surface car ils seront calculés en même temps que le processus de création du *Max-Tree*. Nous allons également étudier des critères couleur d'une image associée à l'image d'entrée sur laquelle nous calculons le *Max-Tree*. Par exemple, des informations couleur comme la moyenne, l'écart-type et le mode, peuvent être estimées au cours de la construction du *Max-Tree*.

### 5.3 Application : Images de façade

Dans la section 4.5 nous avons présenté la segmentation des images de façade en utilisant l'ouverture ultime  $\theta$  et la version qui combine l'information de forme  $\theta^\Omega$ . Nous avons également présenté les raisons pour lesquelles les images de façade peuvent être fortement affectées par les problèmes de masquage en utilisant l'opérateur  $\theta$ , en produisant une sous-segmentation. Les régions de la sous-segmentation  $q_\theta$  ont certaines caractéristiques qui pourraient être exploitées comme critères hiérarchiques. Par exemple, des régions de grande surface et/ou des régions très hétérogènes en termes de couleur sont des régions candidates à être subdivisées. Ainsi, pour utiliser ces caractéristiques, nous proposons deux critères hiérarchiques :

- **Surface** :  $T = Surface(C_h^k) > Seuil_{Surface}$
- **Hétérogénéité Couleur** : l'hétérogénéité sera mesurée à l'aide de l'écart-type de chaque région sur les composantes *RVB*. Donc  $T = \sigma_R(C_h^k) + \sigma_V(C_h^k) + \sigma_B(C_h^k) > Seuil_{\sigma_{\{R,V,B\}}}$

La Figure 5.5 illustre le résultat de l'ouverture ultime par hauteur de l'image gradient inversée de la Figure 5.5(a). Comme nous pouvons l'observer, l'image présente les problèmes de masquage et de fuites, car certaines structures ont été masquées et l'arbre est fusionné au bâtiment.

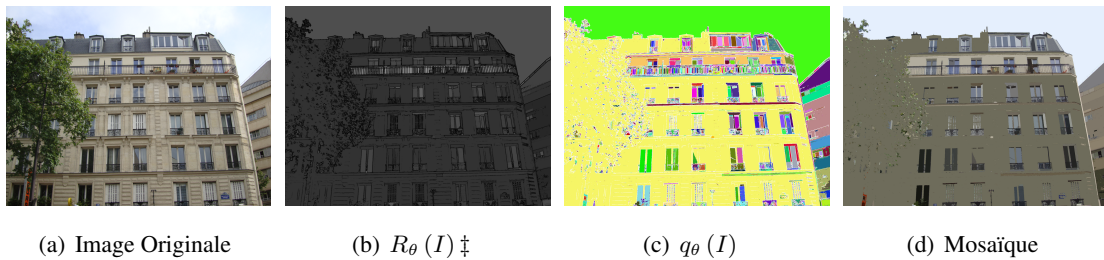


FIGURE 5.5 – (a) Image Originale, (b)-(c)  $R_\theta(I)$  et  $q_\theta(I)$  issues de l'ouverture ultime, (d) Image Mosaïque sur  $q_\theta(I)$ .  $\ddagger$  Correction gamma de 2.



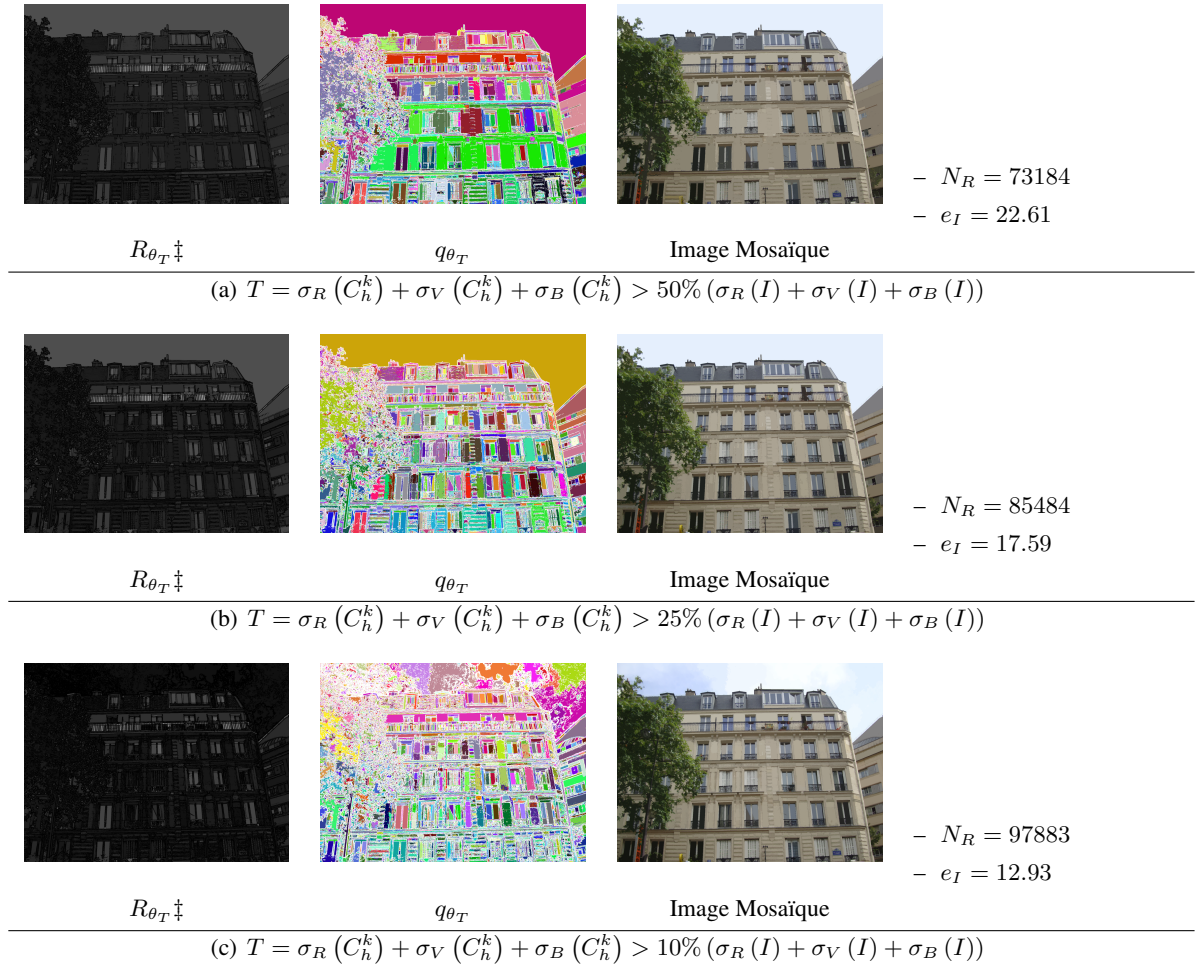
Nous proposons de sélectionner les seuils en fonction de chaque image comme un facteur de l'attribut utilisé ; par exemple, dans le cas de la surface le seuil est fonction de la surface de l'image. Ainsi, des résultats de l'opérateur  $\theta_T$  avec plusieurs seuils pour le critère d'hétérogénéité couleur sont présentés dans la Figure 5.6. De la même façon, nous avons réalisé la même série de tests pour le critère de surface, en donnant comme résultat les images de la Figure 5.7. Les seuils sont un facteur de 50%, 25% et 10% de l'attribut. Par ailleurs, nous allons aussi analyser le nombre de régions de la segmentation et une mesure d'erreur. La mesure d'erreur  $e_I$  correspond à la somme des distances des pixels de l'image originale et des pixels de l'image mosaïque sur la surface de l'image comme l'illustre l'équation 5.11.

$$e_I = \frac{\sum_{i=1}^{N_R} \sum_{p \in R_i} \sum_{x \in \{R, V, B\}} \left( C_x(p) - \hat{C}_x(R_i) \right)^2}{A_I} \quad (5.11)$$

où,  $A_I$  est la surface de l'image,  $N_R$  est le nombre de régions,  $C_x(p)$  est la valeur de la composante  $x$  sur le pixel  $p$ ,  $\hat{C}_x(R_i)$  est la valeur moyenne de la composante  $x$  dans la région  $R_i$ , donc  $\hat{C}_x(R_i) = \frac{\sum_{p \in R_i} C_x(p)}{A_i}$ .

Ainsi, dans le cas de l'ouverture ultime classique, la segmentation présente 25328 régions et 51.41 d'erreur (Dimensions de l'image  $2112 \times 1584$ ).

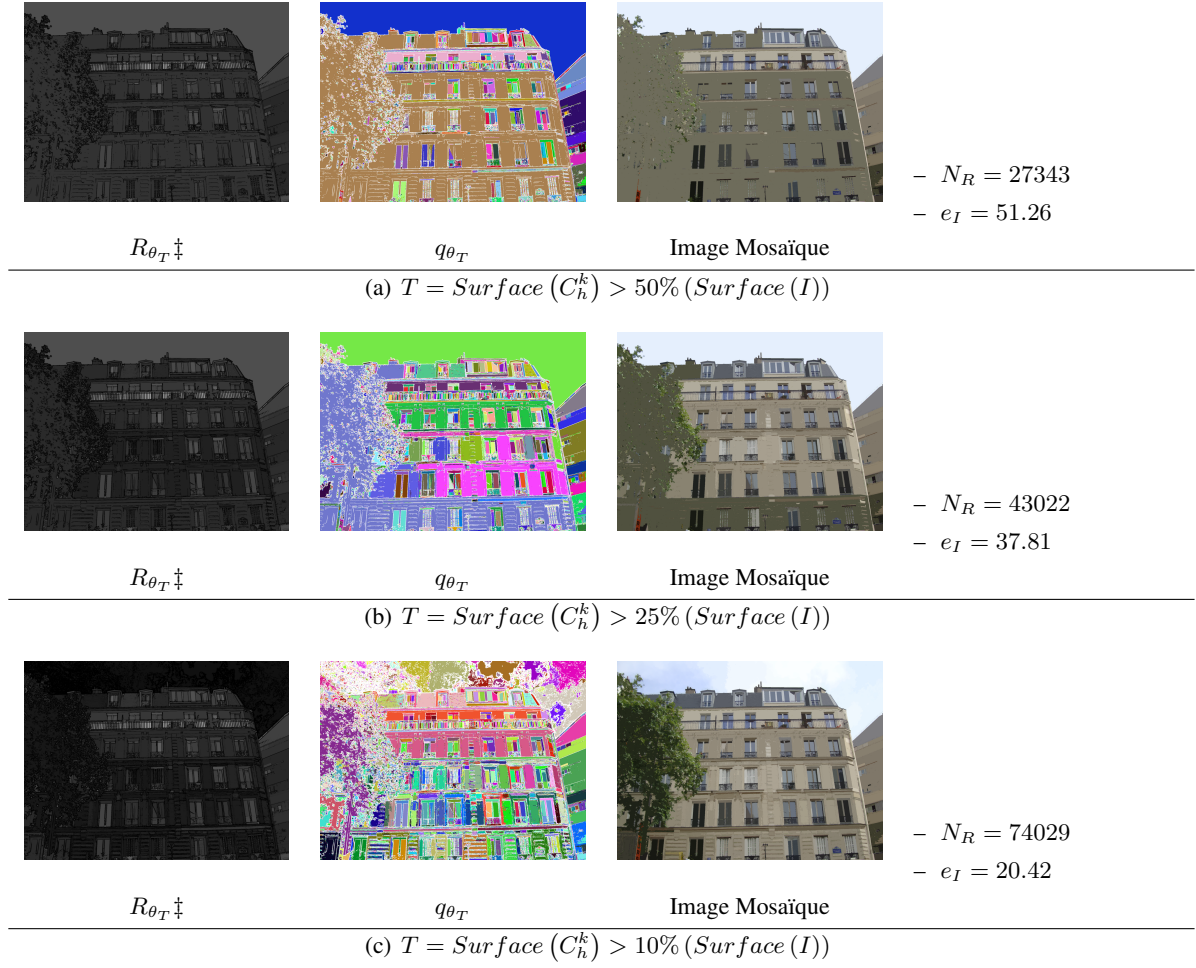
Comme nous avons pu l'observer sur les images, les problèmes ont été réduits sauf dans le cas de 50% pour le critère de surface. Pour ce cas, le nombre de régions est légèrement supérieur à celui que dans le cas de l'ouverture ultime classique, ainsi l'erreur est aussi légèrement inférieure. En utilisant un facteur de 50% du critère couleur et 10% du critère de surface nous pouvons nous apercevoir que le nombre de régions et l'erreur sont approximativement similaires. Néanmoins, en regardant l'image de la segmentation, le ciel est en une seule région pour le cas du critère couleur et divisé en plusieurs régions pour le cas du critère de surface. De manière générale, les résultats correspondent à nos attentes car en descendant la valeur du facteur, le critère est de plus en plus restreint et des régions de plus en plus petites apparaissent. Un comportement que nous pouvons aussi mentionner est le fait que le nombre de régions augmente et la mesure d'erreur diminue. Ainsi, pour une application donnée une bonne segmentation est celle qui met en équilibre ces deux valeurs : nombre de régions et erreur. Nous allons étudier de manière plus approfondie des méthodes d'évaluation non supervisée de la qualité d'une segmentation dans le Chapitre 7, afin de trouver le bon facteur pour les deux critères utilisés. Des résultats sur d'autres images de façade et des comparaisons avec la méthode de l'ouverture ultime avec l'information de forme seront aussi présentés dans ce même chapitre.


 FIGURE 5.6 – (a)-(b)  $R_\theta(I)$  et  $q_\theta(I)$  issues de l’ouverture ultime, (c) Image Mosaïque sur  $q_\theta(I)$ . Correction gamma de 2.

## 5.4 Discussion

Dans ce chapitre nous avons présenté la hiérarchisation de l’ouverture ultime  $\theta_T^N(I)$ , pour un niveau de hiérarchie  $N$  et un critère hiérarchique  $T$ . Pour cela, nous avons d’abord introduit l’opérateur de masquage  $\mathcal{M}_X(I)$  et l’ouverture ultime géodésique  $\theta_X(I)$ . Le nouvel opérateur  $\theta_T^N(I)$  élimine les problèmes si les régions qui produisent les problèmes vérifient le critère hiérarchique choisi.

Même si nous avons défini l’ouverture ultime hiérarchique de manière générale, nous n’avons que mis en œuvre l’opérateur avec des ouvertures par attributs et en utilisant seulement le paramètre critère  $T$ , car ces ouvertures sont facilement calculées à l’aide d’une représentation *Max-Tree* de l’image. Concernant les critères, nous pouvons utiliser tous les attributs géométriques et colorimétriques des nœuds du *Max-Tree* ; cependant pour garder un


 FIGURE 5.7 – Résultats de l'opérateur  $\theta_T$  .

temps de calcul raisonnable, nous nous sommes focalisés sur des attributs simples. Par ailleurs avec l'utilisation d'un critère et en appliquant l'opérateur jusqu'à l'idempotence, nous pouvons descendre dans différents niveaux de hiérarchie de façon indépendante. Ces niveaux, nous les avons nommés niveaux optimaux de la hiérarchie.

Finalement, nous avons présenté l'utilisation de l'opérateur  $\theta_T^N(I)$  sur des images de façade en réduisant les problèmes de masquage. Nous avons utilisé deux critères basés sur des attributs de surface et d'hétérogénéité couleur. L'attribut d'hétérogénéité couleur est basé sur l'écart-type de chaque composante  $RVB$  sur les nœuds du *Max-Tree*. Les seuils utilisés dans la définition du critère sont choisis comme fonction de l'attribut de l'image totale.

Nous avons rapidement utilisé une mesure d'erreur pour comparer l'image segmentée avec l'image originale. Ainsi, en utilisant le critère d'hétérogénéité couleur nous obtenons de meilleurs résultats par rapport au critère de surface.

Ce comportement est logique car la mesure d'erreur est basée sur la distance euclidienne couleur. Cependant, une analyse des mesures d'évaluation non supervisée des méthodes de segmentation sera présentée au Chapitre 7.

# Ouverture Ultime par $\Delta$ –attributs

## Sommaire

<b>6.1</b>	<b>Introduction</b>	<b>73</b>
<b>6.2</b>	<b>Ouverture Ultime par <math>\Delta</math>–attribut</b>	<b>74</b>
6.2.1	Mise en œuvre sur le <i>Max-Tree</i>	79
<b>6.3</b>	<b>Application : Images de texte enfoui</b>	<b>80</b>
6.3.1	Filtrage des valeurs perchées	84
<b>6.4</b>	<b>Discussion</b>	<b>85</b>

## 6.1 Introduction

Retornaz introduit dans sa thèse [129] l’ouverture ultime avec accumulation. Elle consiste à accumuler une série de résidus différents de zéro, engendrés par des ouvertures consécutives. L’objectif de l’accumulation est d’améliorer l’estimation de contraste (transformée  $R_\theta$ ) des objets flous. Retornaz propose également une mise en œuvre basée sur un algorithme d’inondation. Nous verrons que cette implantation peut donner une mauvaise estimation de taille (indicatrice  $q_\theta$ ) à certaines composantes. L’ouverture ultime par  $\Delta$ –attributs, que nous décrivons dans ce chapitre, reprend la notion de l’ouverture ultime par accumulation, où les résidus sont accumulés tant qu’une série d’ouvertures consécutives ne produit pas  $\Delta - 1$  résidus nuls. Nous proposons également une mise en œuvre basée sur le *Max-Tree* qui corrige le problème de mauvaise évaluation de l’indicatrice mentionné précédemment.

Le chapitre est divisé comme suit : la section 6.2 décrit la méthode et présente la mise en œuvre sur le *Max-Tree*, la section 6.3 illustre une application de segmentation de texte enfoui, et finalement la section 6.4 discute de l’efficacité de l’approche .

## 6.2 Ouverture Ultime par $\Delta$ –attribut

L'ouverture ultime par  $\Delta$ –attribut se présente comme une vraie alternative pour combattre le problème dû aux transitions graduelles (images floues). Une structure dont le contour a une transition graduelle voit son contraste sous-estimé. En effet, dans une transition graduelle, un fort résidu est divisé en plusieurs résidus plus petits. Ainsi, le contraste (transformée  $R_\theta$ ) associé à la structure sera le plus grand de ces résidus, par définition plus petit que la transition globale. De cette façon, la probabilité d'être masquée par une structure la contenant augmente. L'objectif de cet opérateur est de détecter ces transitions graduelles et d'accumuler les résidus afin de renforcer le contraste associé à la structure, et donc d'éviter un possible masquage. Une transition est dite graduelle, si des ouvertures de taille  $\lambda_1$  et  $\lambda_2$  avec  $|\lambda_1 - \lambda_2| \leq \Delta$  produisent un résidu non nul.

Observons le profil de la Figure 6.1(a). Nous pouvons identifier plusieurs transitions graduelles dans les pixels  $\{d, e, f\}$ . En calculant les résidus produits par des ouvertures consécutives, les pixels  $\{a, b, c\}$  voient la série de résidus suivants :  $r_1 = 0, r_2 = 0, r_3 = 1, r_4 = 1, r_5 = 2, r_6 = 2, r_7 = 0$  et  $r_8 = 3$ . La Figure 6.1(d) illustre les résidus non nuls de tous les pixels. Ainsi, en utilisant l'ouverture ultime, seul le résidu de valeur 3 est préservé (voir les Figures 6.1(b) et 6.1(c)).

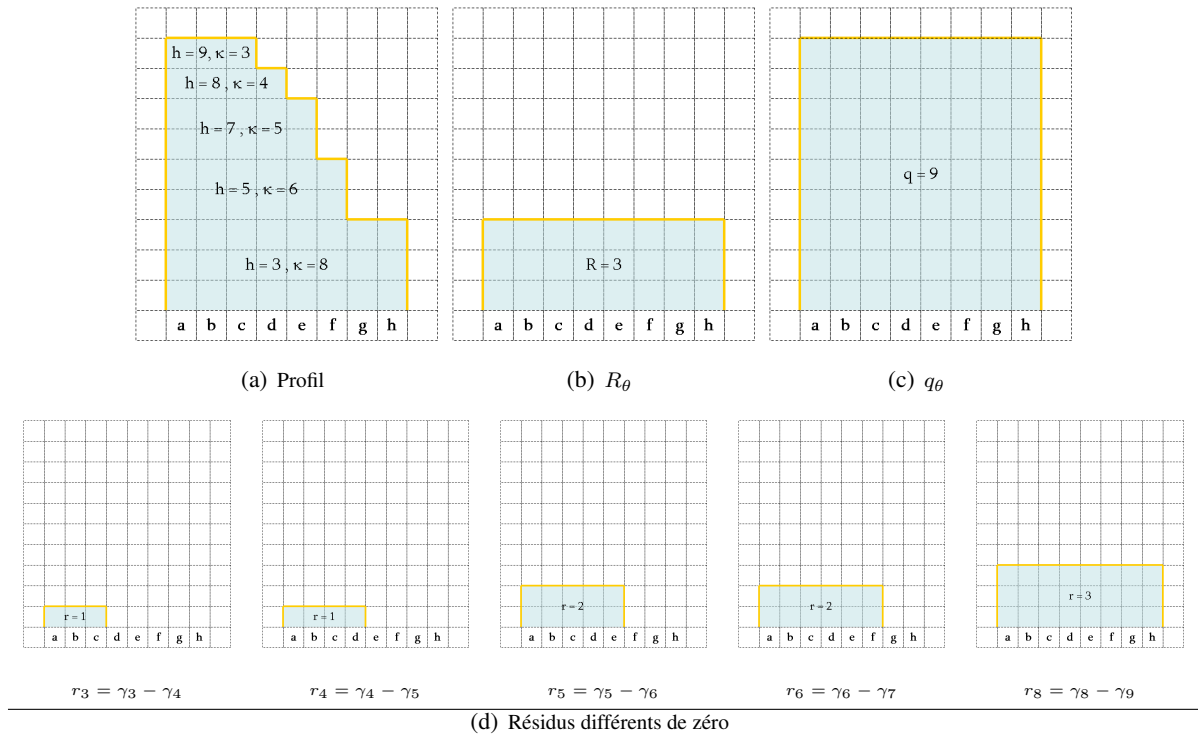


FIGURE 6.1 – Exemple d'un profil avec des transitions graduelles.

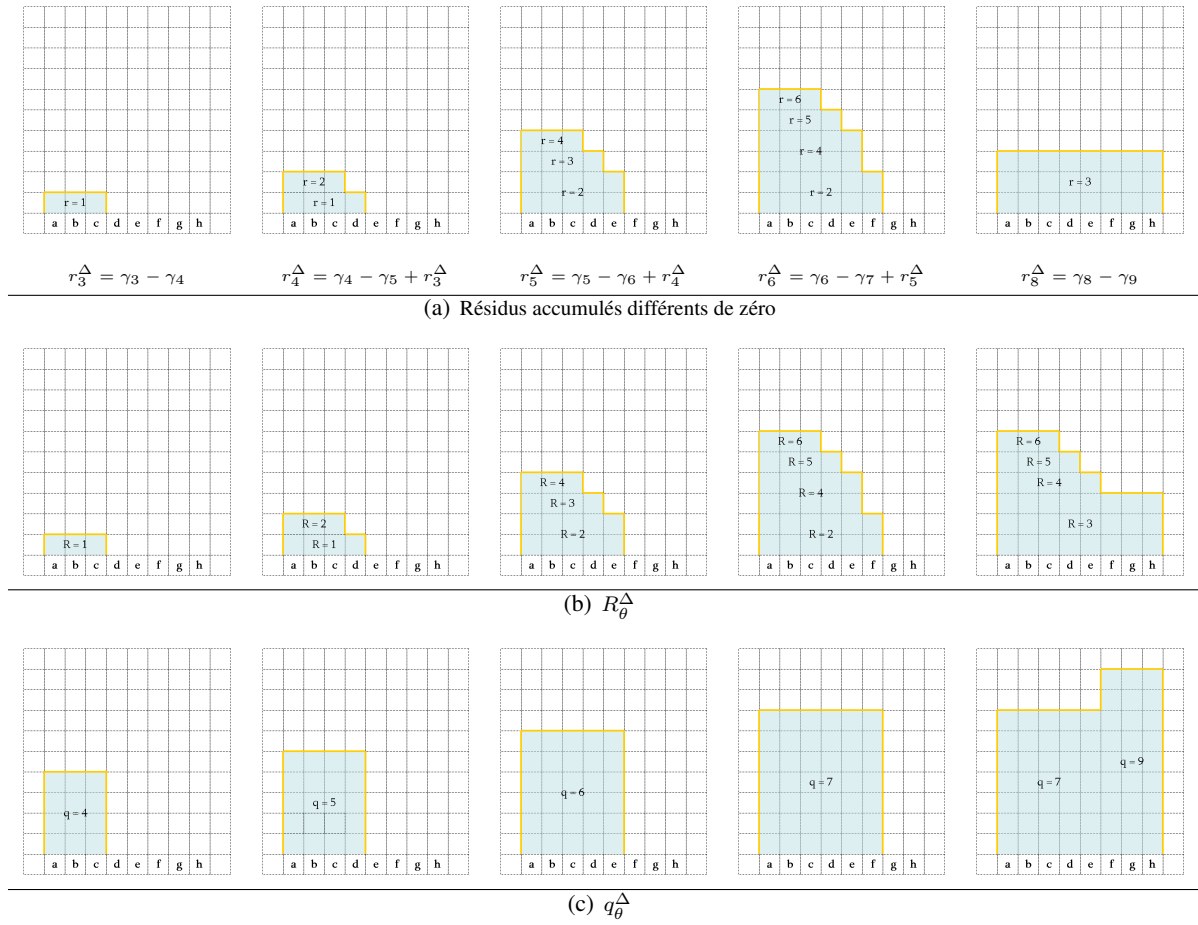
L'ouverture ultime par  $\Delta$ –attribut propose l'accumulation des deux résidus non nuls si la différence de leurs

attributs est inférieure ou égale au paramètre  $\Delta$ . Dans notre exemple, pour une valeur de  $\Delta = 1$  les résidus de  $r_3$  à  $r_6$  seront accumulés ( $r_3 + r_4 + r_5 + r_6 = 6$ ).  $r_8$  ne sera pas accumulé car  $8 - 6 = 2 > \Delta = 1$ . Par contre, pour une valeur de  $\Delta = 2$ ,  $r_8$  sera pris en compte par l'accumulation et  $\sum_{\lambda=3}^8 r_\lambda = 9$ .

Le fait d'accumuler implique que le calcul de résidus doit être réalisé de manière successive, en considérant les ouvertures en ordre croissant ou décroissant (de l'ouverture la plus petite vers la plus grande ou vice-versa). Cette condition n'était pas nécessaire pour la version classique de l'opérateur. Comme nous pouvons l'observer dans les deux parcours de calcul de l'ouverture ultime par  $\Delta$ -attribut (voir les Figures 6.2(b) et 6.3(b)), même si les résidus intermédiaires sont différents, le résultat final de la transformée est toujours le même. Ce résultat est expliqué du fait que l'accumulation des résidus d'une composante peut être vue comme le résidu entre la plus petite ouverture et la plus grande ouverture qui ont contribué à l'accumulation, et finalement la transformée ne sera pas affectée par l'ordre croissant ou décroissant de parcours. Néanmoins, cela n'est pas le cas pour l'indicatrice. Par définition l'indicatrice est la valeur de l'ouverture qui a produit un résidu maximal ; cependant dans l'accumulation plusieurs ouvertures ont participé à la génération d'un résidu maximal.

Voyons dans un premier temps le calcul de l'ouverture ultime avec des résidus accumulés en suivant l'ordre croissant pour une valeur de  $\Delta = 1$ , en mettant à jour l'indicatrice avec la taille de l'ouverture plus un, au fur et à mesure qu'un résidu maximal est produit. Le lecteur peut s'appuyer sur la Figure 6.1(d) pour les résidus classiques ( $r_\lambda$ ) et la Figure 6.2(a) pour les résidus accumulés ( $r_\lambda^\Delta$ ).

1.  $r_3^\Delta = r_3$  car il est le premier résidu généré différent de zéro. Ainsi,  $R_\theta^\Delta = r_3^\Delta$  et  $q_\theta^\Delta = 4$  pour les pixels  $\{a, b, c\}$ .
2.  $r_4 \neq 0$  et donc  $r_4^\Delta = r_4 + r_3^\Delta$ . La transformée et l'indicatrice sont mises à jour car  $r_4^\Delta > r_3^\Delta$ .  $R_\theta^\Delta = r_4^\Delta$  et  $q_\theta^\Delta = 5$  pour les pixels  $\{a, b, c, d\}$ .
3. La procédure d'accumulation continue jusqu'à ce qu'un résidu soit égal à zéro ( $r_7 = 0$ ). Comme nous l'avons mentionné, nous pouvons interpréter cette accumulation de résidus comme le résidu généré entre la plus petite ouverture et la plus grande ouverture ayant contribué à l'accumulation :  $r_6^\Delta = r_6 + r_5^\Delta = r_6 + r_5 + r_4^\Delta = r_6 + r_5 + r_4 + r_3^\Delta = \gamma_3 - \gamma_7$ , et donc  $R_\theta^\Delta = r_6^\Delta$  et  $q_\theta^\Delta = 7$ . A ce moment la composante  $\{a - f\}$  de taille 6 a une bonne évaluation de 7 dans l'indicatrice, ce qui est conforme au comportement de l'ouverture ultime.
4. Le résidu suivant différent de zéro est  $r_8^\Delta = r_8$ . Bien que le pixel  $\{f\}$  ait participé à l'accumulation précédente, il est modifié par le résidu  $r_8^\Delta$ , car  $r_8^\Delta\{f\} > r_6^\Delta\{f\}$ . Par contre,  $r_8^\Delta$  ne changera pas l'évaluation des pixels  $\{a - e\}$  car leur transformée courante est supérieure à  $r_8^\Delta\{a - e\}$ .  $R_\theta^\Delta$  et  $q_\theta^\Delta$  sont mises à jour seulement pour les pixels  $\{f, g, h\}$ . Le changement d'évaluation de  $q_\theta^\Delta\{f\}$  "rogne" d'un pixel la composante dont la taille avait été évaluée à 7, en produisant une différence de taille de la composante connexe et de son évaluation d'indicatrice. Le problème de mauvaise évaluation de l'indicatrice a été étudié par


 FIGURE 6.2 – Calcul de l'ouverture ultime par  $\Delta$ –attribut depuis l'ouverture la plus petite jusqu'à la plus grande.

Beucher [14] comme des valeurs perchées de l'ouverture ultime.

Voyons maintenant le calcul de l'ouverture ultime avec des résidus accumulés depuis l'ouverture la plus grande vers la plus petite, en mettant à jour l'indicatrice avec la taille de l'ouverture plus un, au fur et à mesure qu'un résidu maximal est généré (voir Figure 6.3).

1. Le premier résidu généré différent de zéro est  $r_8^\Delta = r_8$ .  $R_\theta^\Delta = r_8^\Delta$  et  $q_\theta^\Delta = 9$ .
2. Le résidu suivant  $r_7 = 0$  et nous n'accumulons pas.
3.  $r_6^\Delta = r_6$  et  $R_\theta^\Delta$  et  $q_\theta^\Delta$  ne changent pas car  $r_6^\Delta < r_8^\Delta$ .
4.  $r_5 \neq 0$  alors  $r_5^\Delta = r_5 + r_6^\Delta$ . Ce résidu produit une mise à jour de  $R_\theta^\Delta = r_5^\Delta$  et  $q_\theta^\Delta = 6$  pour les pixels  $\{a - e\}$ .
5. A partir de ce moment chaque résidu accumulé produira un résidu maximal de la manière suivante :
  - $r_4^\Delta > r_5^\Delta$ , donc  $R_\theta^\Delta = r_4^\Delta$  et  $q_\theta^\Delta = 5$  pour les pixels  $\{a, b, c, d\}$ ,




 FIGURE 6.3 – Calcul de l'ouverture ultime par  $\Delta$ -attribut depuis l'ouverture la plus grande jusqu'à la plus petite.

–  $r_3^\Delta > r_4^\Delta$ , donc  $R_\theta^\Delta = r_3^\Delta$  et  $q_\theta^\Delta = 4$  pour les pixels  $\{a, b, c\}$ .

En résumé, si nous mettons à jour l’indicatrice au fur et à mesure qu’un résidu important est généré, nous avons deux comportements différents au niveau de l’indicatrice dépendant du type de parcours :

- Dans un calcul croissant, la valeur de l’indicatrice sera la taille de la dernière ouverture qui a contribué à l’accumulation. Ce résultat produit une seule composante au niveau de l’indicatrice, mais dans certains cas elle sera surévaluée car d’autres résidus maximaux peuvent “rogné” la composante accumulée (comme nous avons vu pour le cas du pixel  $\{f\}$  dans l’illustration de la Figure 6.2(c)).
- Dans un calcul décroissant, la valeur de l’indicatrice changera au fur et à mesure que les résidus sont accumulés à partir du moment où un résidu accumulé dépasse la valeur courante de la transformée. Dans notre exemple  $r_6^\Delta > r_8^\Delta$  et  $q_\theta^\Delta\{e\} = 6$ ,  $q_\theta^\Delta\{d\} = 5$  et donc  $q_\theta^\Delta\{a, b, c\} = 4$ . Ce résultat n’est pas intéressant car l’objectif de l’ouverture ultime avec accumulation est aussi d’assigner la même valeur d’indicatrice aux résidus accumulés.

Afin de réduire le problème des valeurs perchées, nous proposons de bloquer la valeur de l’indicatrice, dans le sens décroissant de calcul, à partir du moment où un résidu accumulé dépasse la valeur courante (intermédiaire) de la transformée. Dans notre exemple quand on va estimer la transformée du pixel  $\{f\}$ , on sait que  $r_6^\Delta < r_8^\Delta$  et le pixel  $f$  ne contribue pas à l’estimation de la taille de la composante (indicatrice), même si le résidu  $r_6^\Delta$  participe à l’accumulation. Par contre,  $r_5^\Delta$  dépasse la valeur courante de la transformée et donc la valeur de l’indicatrice de  $5 + 1 = 6$  sera assignée au reste des pixels participant à l’accumulation  $q_\theta^\Delta\{a - e\} = 6$  (voir Figure 6.3(d)). Cette alternative peut être vue comme la “propagation” de l’indicatrice. Ainsi, les résidus accumulés qui dépassent la valeur courante de la transformée auront une seule valeur d’indicatrice avec la bonne évaluation.

Par ailleurs, nous pouvons définir la transformée de la même manière que dans la définition de l’ouverte ultime classique, c’est-à-dire  $R_\theta$  est égal aux résidus maximaux de la série de résidus avec l’accumulation, comme l’illustre la Définition 17.

**Définition 17.** Soit  $I$  une image à niveaux de gris, pour tout pixel  $x$ , la transformée de l’ouverture ultime par  $\Delta$ -attribut est décrite par :

$$R_\theta(x) = \bigvee (r_\lambda(x)), \quad \forall \lambda \geq 1$$

$$\text{où } r_\lambda(x) = \begin{cases} 0 & \gamma_\lambda(x) = \gamma_{\lambda+\Delta}(x) \\ (\gamma_\lambda(x) - \gamma_{\lambda+\Delta}(x)) + r_{\lambda\pm 1}(x) & \text{sinon} \end{cases} \quad (6.1)$$

$\gamma_\lambda(x)$  est une ouverture de taille  $\lambda$ , et  $r_{\lambda-1}(x)$  et  $r_{\lambda+1}(x)$  pour un calcul croissant et décroissant des ouvertures respectivement.

Le nouvel opérateur peut réduire le problème des transitions graduelles si les transitions sont détectées par des changements d’attribut de taille inférieure ou égale à  $\Delta$ . L’image synthétique de la Figure 6.4(a) montre une

transition en deux pas <sup>1</sup>. Le pas intermédiaire a deux pixels d'épaisseur. Si nous utilisons une valeur de  $\Delta \leq 2$ , le rectangle continuera d'être masqué. Par contre en utilisant une valeur de  $\Delta = 3$  le problème est résolu. C'est pour cela que le choix du paramètre  $\Delta$  reste important pour une application donnée.

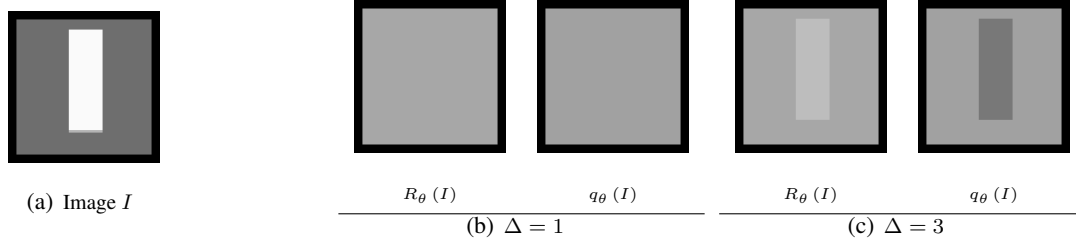


FIGURE 6.4 – Exemple : Transitions graduelles (a) Image Synthétique, (b) et (c) Ouverture ultime par  $\Delta$ -attribut.

### 6.2.1 Mise en œuvre sur le *Max-Tree*

Retornaz a proposé une implantation par inondation de l'opérateur avec accumulation, ainsi cette version calcule la transformée depuis l'ouverture la plus petite jusqu'à la plus grande. Par contre, dans le *Max-Tree* le calcul se réalise dans le sens inverse. C'est pourquoi la proposition de la propagation de l'attribut du nœud, lorsque son résidu dépasse la valeur courante de la transformée, peut être mise en œuvre sur cette structure. Dans la mise en œuvre de la version classique de l'ouverture ultime sur le *Max-Tree*, les résidus sont accumulés si l'attribut d'un nœud  $C_h^k$  est égal à celui de son parent  $C_{h'}^{k'}$ , car le résidu est la différence entre son niveau de gris  $h$  et le niveau de gris de son premier ancêtre avec un attribut différent [41]. Afin d'intégrer le paramètre  $\Delta$  sur le calcul des résidus, nous changeons la condition d'égalité entre les attributs père-fils, à la condition qu'une différence entre eux soit inférieure ou égale à la valeur  $\Delta$ , comme l'illustre l'équation 6.2.

$$r_\lambda(C_h^k) = \begin{cases} h - h' + r_\lambda(C_{h'}^{k'}) & \kappa_{C_h^k} - \kappa_{C_{h'}^{k'}} \leq \Delta \\ h - h' & \text{sinon} \end{cases} \quad (6.2)$$

La Figure 6.5 montre le *Max-Tree* du profil de la Figure 6.1(a) et le calcul des résidus de la version classique de l'ouverture ultime et de la version avec l'accumulation. Ainsi, le résidu peut être vu comme la différence entre le niveau de gris du nœud et le niveau de gris du premier nœud qui a participé à l'accumulation.

Comme nous l'avons mentionné, le calcul de  $R_\theta$  se réalise de la même manière que pour la version classique, où chaque enfant compare son résidu avec le résidu maximal de son parent  $R_\theta(C_{h'}^{k'})$ , et  $R_\theta(C_h^k)$  enregistre la valeur maximale entre eux. Pour le calcul de  $q_\theta(C_h^k)$ , nous propageons de père à fils l'attribut plus un du nœud à

1. Le lecteur peut se référer à la page 43 et la Figure 3.7 pour observer le problème dû aux transitions graduelles sur cette image synthétique.

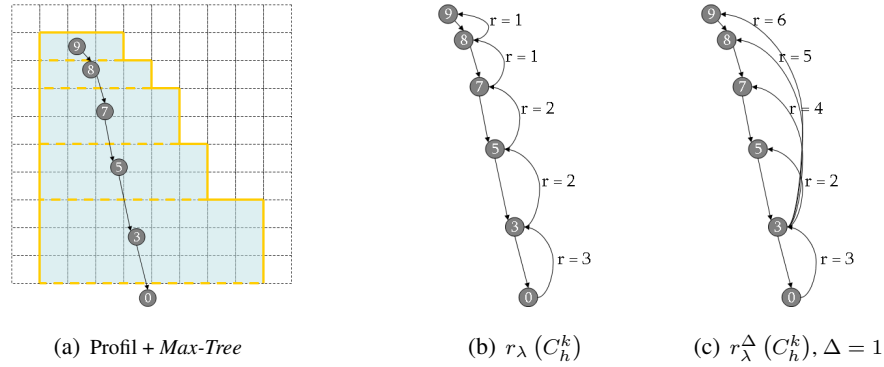


FIGURE 6.5 – (a) Max-Tree d’un profil avec des transitions graduelles. (b) Calcul de résidus classique. (c)- (d) Calcul de l’accumulation des résidus avec  $\Delta = 1$

partir du moment où son résidu dépasse la valeur courante de la transformée dans une série d’accumulations. Par exemple pour le profil de la Figure 6.5, nous assignons l’attribut plus un de nœud  $C_7^0$  à tous ces enfants lorsque l’accumulation est activée.

### 6.3 Application : Images de texte enfoui

Le texte enfoui (*scene-text*) est lié à la sémantique de l’image et constitue une information pertinente pour l’indexation d’une image basée sur son contenu [130]. Cette recherche est développée dans le cadre du projet ITOWNS (*Image-based Town On-line Web navigation and Search Engine*). L’objectif de ce projet est de mettre au point une nouvelle génération d’outils multimédia sur le web qui combine un navigateur 3D géographique (comme le Geoportail, Google Earth, Microsoft live Earth) avec un moteur de recherche basé sur une indexation par le contenu des données images/visuelles.

Dans la plupart des cas, le texte enfoui sur les images est placé sur des surfaces de différentes couleurs, telles que pancartes, affiches, etc. . . qui favorisent la visibilité des lettres. Mais ces surfaces sont également contrastées avec leur environnement. Lorsque nous utilisons l’ouverture ultime, les caractères peuvent être masqués parce que le contraste entre la surface et son environnement est supérieur au contraste entre les lettres et la surface. Par ailleurs, à cause de la numérisation, la bordure des lettres peut être floue en produisant des transitions graduelles.

Un exemple de l’ouverture ultime par  $\Delta$ – attribut est illustré sur la Figure 6.6. L’image contient le texte “PAS-SPORT». En utilisant l’ouverture ultime par attribut de hauteur, les lettres “O» et “R» ont été masquées. Comme nous pouvons l’observer, en utilisant la version avec l’accumulation, toutes les lettres sont détectées et le contraste des autres lettres a presque doublé. Par exemple, le “A» est passé de 55 à 86 et le “S» de 48 à 98.

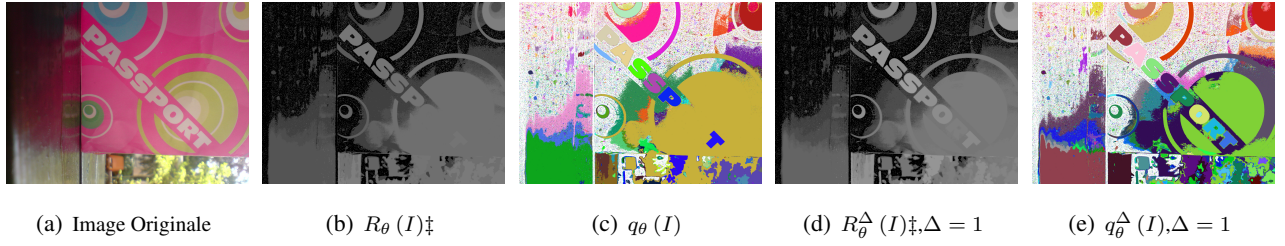


FIGURE 6.6 – (a) Image Originale. (b)- (c) Ouverture ultime classique. (d)- (e) Ouverture ultime avec  $\Delta$ -attribut  $\Delta = 1$ .  $\ddagger$  Correction gamma de 2.

Un autre exemple de l'utilisation de l'ouverture ultime par  $\Delta$ -attribut est présenté dans la Figure 6.7 avec  $\Delta = 1$  et  $\Delta = 2$ . Au niveau de la transformée, nous pouvons apprécier une augmentation considérable du contraste des lettres. Dans le cas particulier de  $\Delta = 1$  seules certaines lettres, comme le “i» de “adhesive» et le “a» et “s» de “address», n'ont pas un contraste aussi fort que les autres. Ce comportement est produit car ces lettres ont des différences d'attribut supérieures à 1 dans la transition graduelle. Lorsque nous appliquons l'opérateur avec  $\Delta = 2$ , nous pouvons observer une transformée plus homogène. Néanmoins, au niveau de l'indicatrice l'opérateur produit quelques résultats indésirables car des lettres ont été fusionnées, par exemple le “dh» de “adhesive» et le “dd» de “address».

Nous illustrons aussi l'évolution de deux pixels sur les lettres “l» de “self» et “i» de “adhesive» dans la Figure 6.8. Nous pouvons aussi apprécier les résidus qui sont accumulés avec  $\Delta = 1$  et  $\Delta = 2$ . Comme nous l'avons mentionné dans le paragraphe précédent, les évolutions corroborent le fait que la contribution de l'accumulation de résidus est plus importante dans la lettre “l» que dans la lettre “i» pour une valeur de  $\Delta = 1$ , car la valeur de la transformée du pixel sur le “l» est égale à 123 et celle du pixel sur le “i» est de 58. Par contre leurs augmentations sont similaires, pour une valeur de  $\Delta = 2$ . Pour le cas du pixel sur le “i», nous pouvons confirmer avec  $r_\theta(x)$  que la transition de la lettre a plusieurs valeurs résidus de résidus nuls entre  $\lambda = 130$  et  $\lambda = 150$  (voir Figure 6.8(b)). Ceci explique la différence de l'accumulation entre les résidus de  $\Delta = 1$  et de  $\Delta = 2$ .

Même si nous avons illustré dans le premier exemple la réduction du problème de masquage, l'accumulation de résidus peut présenter des résultats avec un masquage encore plus fort. De la même façon que les lettres sont favorisées par l'accumulation, les structures qui les contiennent peuvent l'être aussi, si elles sont floues également. Un exemple de cet effet est illustré sur la Figure 6.9. En utilisant l'opérateur classique, seule une lettre a été masquée, alors qu'avec l'opérateur avec accumulation, le problème du masquage est accentué.



FIGURE 6.7 — Exemple : Transitions graduelles (a) Image Synthétique, (b) et (c) Ouverture ultime par  $\Delta$ —attribut. ‡ Correction gamma de 2.

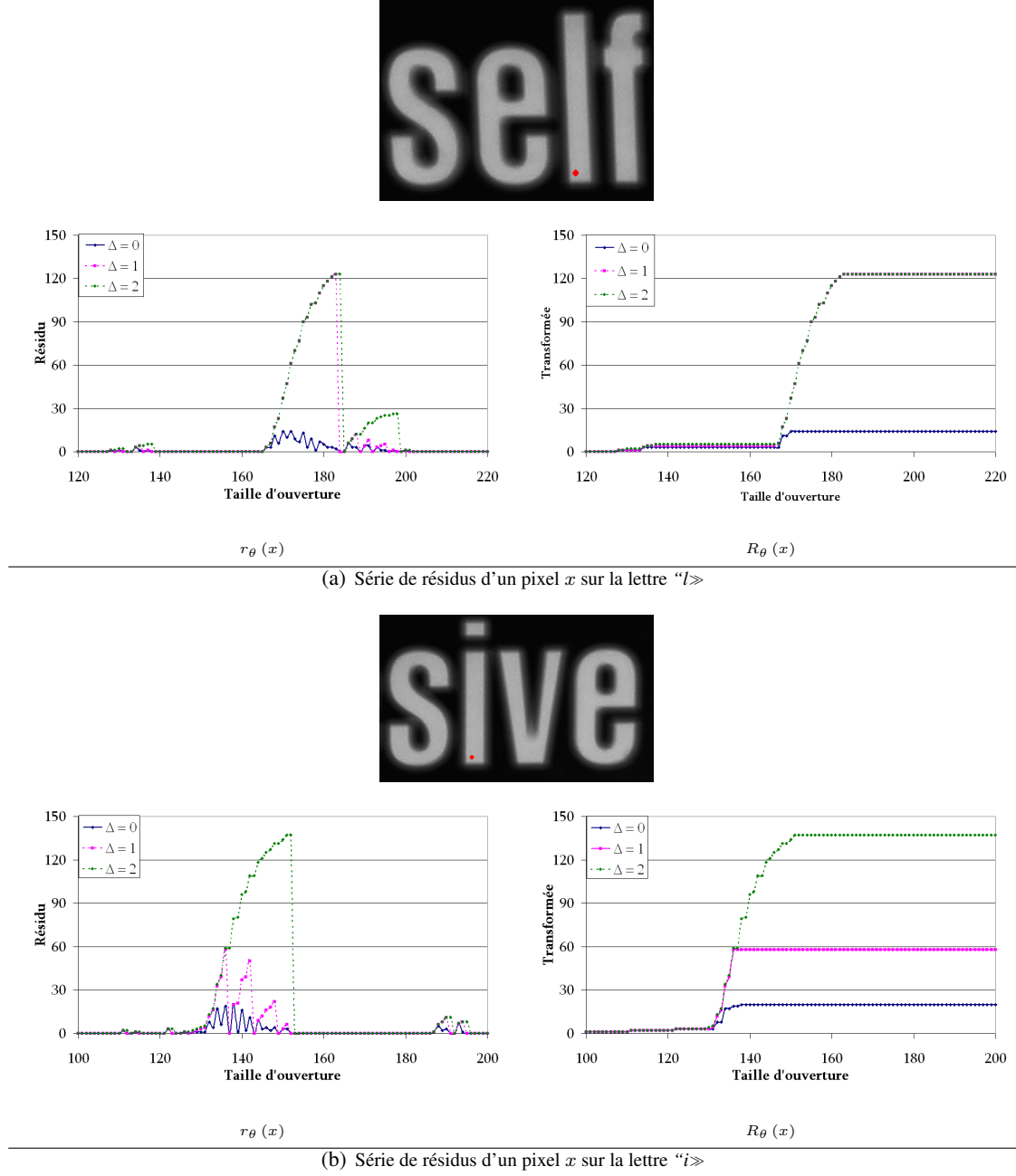


FIGURE 6.8 – Série de résidus de deux pixels.

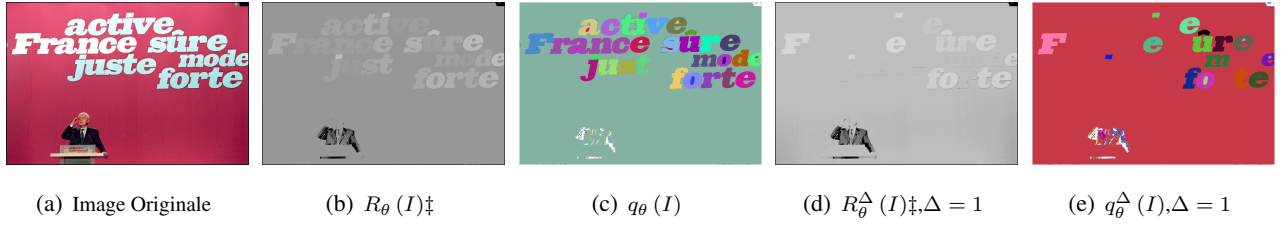


FIGURE 6.9 – Problème par des structures imbriquées. (a) Image Originale. (b)- (c) Ouverture ultime classique. (d)- (e) Ouverture ultime avec  $\Delta$ -attribut  $\Delta = 1$ .  $\ddagger$  Correction gamma de 2.

### 6.3.1 Filtrage des valeurs perchées

Dans une chaîne de reconnaissance de texte, Retornaz a proposé un filtrage des composantes dont l'évaluation de l'indicatrice ne correspond pas à la taille réelle (composantes mal évaluées). Ce filtrage a été appliqué à la sortie de la version classique de l'ouverture ultime, mais il n'a pas pu être étendu à la chaîne de reconnaissance qui utilise l'accumulation, car la mise en œuvre par inondation produit des composantes de l'indicatrice mal évaluées, tel que nous l'avons décrit dans la section précédente. La Figure 6.10 illustre le filtrage de valeurs perchées en utilisant l'opérateur avec l'accumulation par inondation et en utilisant la version *Max-Tree*. Comme nous l'avons mentionné, la mise en œuvre par inondation calcule la version par accumulation depuis l'ouverture la plus petite jusqu'à la plus grande, en produisant des composantes mal évaluées au niveau de l'indicatrice. En éliminant les valeurs perchées, certaines lettres ont été supprimées (voir Figure 6.10(b), les lettres "A, E et M»). En revanche avec la version *Max-Tree*, le problème est résolu.

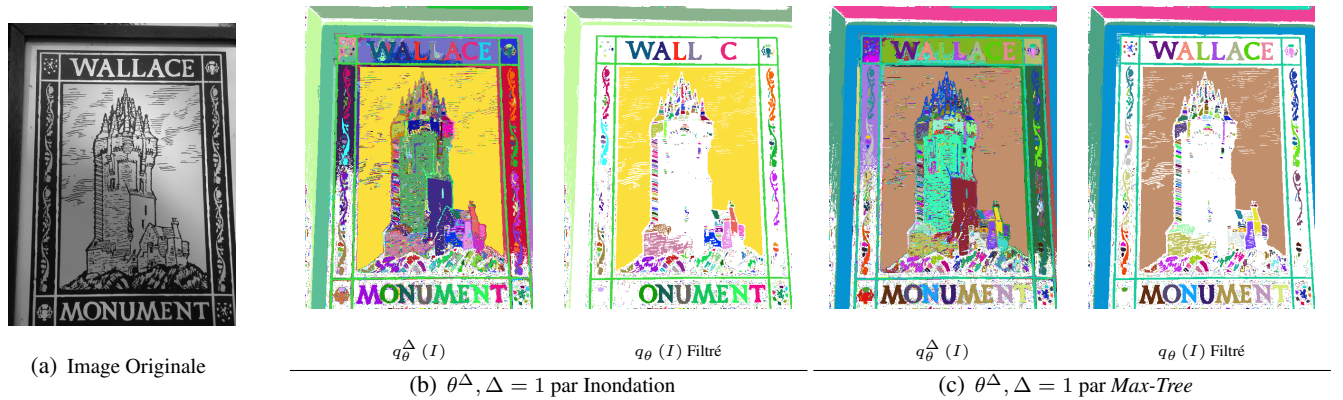


FIGURE 6.10 – Exemple : Filtrage des valeurs perchées (a) Image Originale, Ouverture ultime par  $\Delta$ -attribut par : (b) inondation et (c) *Max-Tree*.



## 6.4 Discussion

Dans ce chapitre nous avons présenté une variante de l'ouverture ultime par attribut qui résout principalement le problème dû aux transitions graduelles. Le nouvel opérateur est basé sur l'accumulation de résidus tant qu'une série d'ouvertures consécutives ne produit pas  $\Delta - 1$  résidus nuls. Nous avons également analysé les différences au niveau de l'indicatrice, lorsque nous calculons l'opérateur depuis l'ouverture la plus petite vers la plus grande et vice-versa. C'est pourquoi nous avons proposé de calculer l'accumulation dans le sens décroissant et en assignant comme indicatrice la valeur de l'ouverture à partir du moment où le résidu accumulé dépasse la valeur de la transformée courante. Nous avons aussi présenté la mise en œuvre de l'opérateur sur le *Max-Tree* en corrigeant le problème présenté dans la version inondation des composantes mal évaluées.

Le principal avantage de l'ouverture ultime par  $\Delta$  – attributs est d'avoir une information plus pertinente par rapport au contraste réel de l'objet car elle prend en compte les accumulations des divers résidus produits par variations d'attributs inférieurs à  $\Delta$ . Cependant, l'accumulation ne corrige pas le problème dû aux structures imbriquées et elle peut même produire des masquages plus importants si la structure qui la contient est floue elle aussi. Par ailleurs, l'accumulation peut augmenter les problèmes de fuites en fusionnant les composantes d'intérêt. Afin d'éviter ces problèmes et de garder les résidus accumulés, nous pouvons imaginer une combinaison de cette approche avec les deux approches présentées antérieurement : information de forme et/ou hiérarchique.

**Sommaire**

---

<b>7.1</b>	<b>Introduction</b>	<b>87</b>
<b>7.2</b>	<b>Segmentation d'images de façade</b>	<b>87</b>
7.2.1	Evaluation de la qualité d'une segmentation	87
7.2.2	Ouverture ultime	91
7.2.3	Ouverture ultime avec l'information de forme	92
7.2.4	Ouverture ultime hiérarchique	93
7.2.5	Ouverture ultime par $\Delta$ -attributs	96
7.2.6	Discussion	97
<b>7.3</b>	<b>Détection de texte enfoui</b>	<b>98</b>
7.3.1	Ouverture ultime avec l'information de forme	100
7.3.2	Ouverture ultime hiérarchique	101
7.3.3	Ouverture ultime par $\Delta$ -attributs	101
7.3.4	Résultats	101
7.3.5	Discussion	103
<b>7.4</b>	<b>Segmentation d'images de cellules</b>	<b>107</b>
7.4.1	Introduction	107
7.4.2	Ouverture ultime avec l'information de forme	107
7.4.3	Résultats	108
7.4.4	Discussion	108
<b>7.5</b>	<b>Discussion Générale</b>	<b>109</b>
7.5.1	Image d'entrée	110
7.5.2	Paramétrage	111

---

## 7.1 Introduction

Dans les chapitres précédents nous avons présenté trois méthodes afin d'améliorer les résultats de segmentation de l'ouverture ultime classique en réduisant les problèmes de masquage et de fuites. Pour démontrer la performance de ces méthodes, dans ce chapitre nous comparons les résultats de différentes approches sur des images issues des trois applications suivantes : analyse de façades, localisation de texte et segmentation de cellules

## 7.2 Segmentation d'images de façade

Nous avons déjà présenté, dans les sections 4.5 et 5.3, l'utilisation de l'ouverture ultime avec l'information de forme et de l'ouverture ultime hiérarchique sur des images de façade. Dans cette section nous comparons les trois approches proposées.

Jusqu'à présent, nous avons estimé les améliorations proposées de l'ouverture ultime par rapport à la version classique à partir d'une comparaison qualitative-visuelle des résultats. Néanmoins, afin de mesurer l'amélioration, nous proposons d'étudier des mesures de qualité qui nous permettront d'analyser les résultats de segmentation de manière quantitative.

### 7.2.1 Evaluation de la qualité d'une segmentation

L'objectif principal d'une mesure d'évaluation est d'être capable de comparer les différents résultats de segmentation et ainsi trouver la meilleure segmentation possible. De nombreuses mesures de segmentation d'images ont été proposées dans la littérature [179]. Ces mesures d'évaluation peuvent être divisées en deux catégories principales : mesures d'évaluation supervisée et mesures d'évaluation non supervisée. Dans notre cas d'étude, nous sommes intéressés par les méthodes non supervisées, étant donné l'absence d'une segmentation de référence (vérité terrain).

#### L'état de l'art

Nous avons condensé dans la Table 7.1 des mesures d'évaluation non supervisée. Liu et Yang [92] ont proposé une mesure prenant en compte le nombre de régions de l'image segmentée ( $N_R$ ), la surface de chaque région ( $A_i$ ), la somme des distances entre la couleur des pixels de chaque région et la couleur correspondante dans l'image segmentée ( $e_i$ ). Dans l'expression d'évaluation, ils pénalisent les sur-segmentations par un facteur dépendant du nombre de régions ( $\sqrt{N_R}$ ) et les petites régions avec une grande différence couleur ( $1/\sqrt{A_i}$ ). Le meilleur résultat de segmentation est celui qui présente la valeur la plus petite. Cette mesure a été conçue pour des segmentations avec peu de régions et des objets bien définis et homogènes dans l'image. Par ailleurs, Borsotti *et al.* [17] ont identifié les limitations de la fonction d'évaluation de Liu et Yang. L'idée principale de Borsotti est de combiner deux

Approche	Expression
Liu et Yang [92]	$F(I) = \frac{\sqrt{N_R}}{1000A_I} \sum_{i=1}^{N_R} \frac{e_i^2}{\sqrt{A_i}}$
Borsotti <i>et al.</i> [17]	$Q(I) = \frac{\sqrt{N_R}}{1000A_I} \sum_{i=1}^{N_R} \left( \frac{e_i^2}{1 + \log A_i} + \left( \frac{N(A_i)}{A_i} \right)^2 \right)$
Glory <i>et al.</i> [54]	$C(I) = \frac{\sum_{i=1}^{N_R} \frac{e_i^2}{1 + \log A_i}}{\frac{e^2}{1 + \log A} + 1} + \frac{\sum_{i=A_{\min}}^{A_{\max}} \frac{N(A_i)}{A_i^2}}{N_R}$
Zhang <i>et al.</i> [178]	$H_l(R_i) = - \sum_{m \in V_i} \frac{L_i(m)}{A_i} \log \frac{L_i(m)}{A_i}$ $H_r(I) = \sum_{i=1}^{N_R} \left( \frac{A_i}{A} \right) H_l(R_i)$ $H_l(I) = - \sum_{i=1}^{N_R} \frac{A_i}{A} \log \frac{A_i}{A}$ $E(I) = H_r(I) + H_l(I)$

où

- $e_i^2 = \sum_{p \in R_i} \sum_{x \in \{R, V, B\}} \left( C_x(p) - \hat{C}_x(R_i) \right)^2$ ,
- $N_R$  nombre de régions,
- $A$  surface de l'image,
- $A_i$  surface de la région  $R_i$ ,
- $\hat{C}_x(R_i) = \frac{\sum_{p \in R_i} C_x(p)}{A_i}$ ,
- $N(A_i)$  nombre de régions de surface  $A_i$ ,
- $V_i$  histogramme de la luminance de la région  $R_i$ ,
- $L_i(m)$  nombre de pixels de luminance  $m$  dans la région  $R_i$ .

TABLE 7.1 – Mesures d'évaluation de résultats de segmentation d'images.

termes : le premier pénalise les régions non homogènes, alors que le second pénalise les régions avec beaucoup de régions de la même taille. Le principal problème de cette mesure est la non-normalisation de ces termes. De plus, dans notre cas d'étude, cette mesure ne remplit pas nos exigences comme méthode d'évaluation. Par exemple dans les images de façade une segmentation peut avoir plusieurs régions de la même surface et elle sera donc pénalisée par le deuxième terme (de fenêtres de la même taille). Une troisième approche, présentée par Glory *et al.* [54], propose une mesure  $C(I)$  de deux termes "normalisés" qui contrôlent l'hétérogénéité couleur de chaque région et la présence de petites régions. Cependant, même s'ils parlent de termes normalisés, la réalité est complètement différente. En effet, pour l'hétérogénéité couleur, la valeur de 1 n'est pas nécessairement la limite supérieure de la métrique, et pour la métrique du nombre de petites régions la valeur de 0 n'est pas la limite inférieure. Les trois approches citées mesurent l'hétérogénéité couleur en utilisant la distance couleur ( $e_i$ ). En revanche, Zhang *et al.* [178] proposent de nouvelles entropies pour l'évaluation du résultat de segmentation d'images. Selon eux, une bonne évaluation de segmentation doit maximiser l'uniformité des pixels dans chaque région segmentée et minimiser l'uniformité entre régions adjacentes. Ils mesurent l'entropie de la luminance de chaque région afin d'estimer l'uniformité d'une région ( $H_r(I)$ ). Lorsque tous les pixels dans une région ont la même valeur, l'entropie pour la région est égale à 0. C'est pourquoi ils utilisent également une entropie du nombre de régions ( $H_l(I)$ ) pour pénaliser les sur-segmentations.

Afin de tester ces métriques, nous avons utilisé la hiérarchisation de l'ouverture ultime avec un critère hiérarchique  $T = Surface(C_h^k) > \beta \times Surface(I)$ . Nous avons fait varier la valeur de  $\beta$ , de 100% jusqu'à 0% chaque 5%. La Figure 7.1 illustre les quatre mesures calculées et normalisées de manière indépendante. Néanmoins, les résultats sont un peu inattendus, car pour toutes les mesures de qualité, le meilleur résultat est celui de  $\beta = 100\%$ , c'est-à-dire l'ouverture ultime classique. Comme nous pouvons l'observer, dans cet exemple l'ouverture ultime présente un fort problème de masquage dans le bâtiment et l'arbre est fusionné aussi avec le bâtiment. Dans l'intervalle  $\beta = \{100, 60\}\%$  les mesures ne changent pas, donc cela correspond à la taille de la région la plus grande de l'image. Par ailleurs, si nous observons localement l'intervalle  $\beta = \{45, 0\}\%$ , nous pourrions établir comme meilleure segmentation celle de  $\beta = 25\%$ . Visuellement le résultat est meilleur que l'ouverture ultime classique. Cependant, l'arbre reste collé au premier étage du bâtiment. Nous avons testé ces mesures avec d'autres images et l'ouverture classique est toujours classée première. Ce comportement est dû au fait que la sortie de l'ouverture ultime hiérarchique présente de petites régions très contrastées, faisant diminuer quantitativement la qualité de la segmentation.

Après avoir analysé les différentes approches de la littérature, nous pouvons conclure que la définition d'une évaluation de la qualité d'une segmentation n'est pas une tâche triviale, qui dépend de l'application et des méthodes à comparer. C'est pourquoi nous continuerons à comparer visuellement les approches, et à utiliser l'information de l'hétérogénéité couleur et du nombre de régions séparément comme de simples indicateurs pour évaluer la qualité.

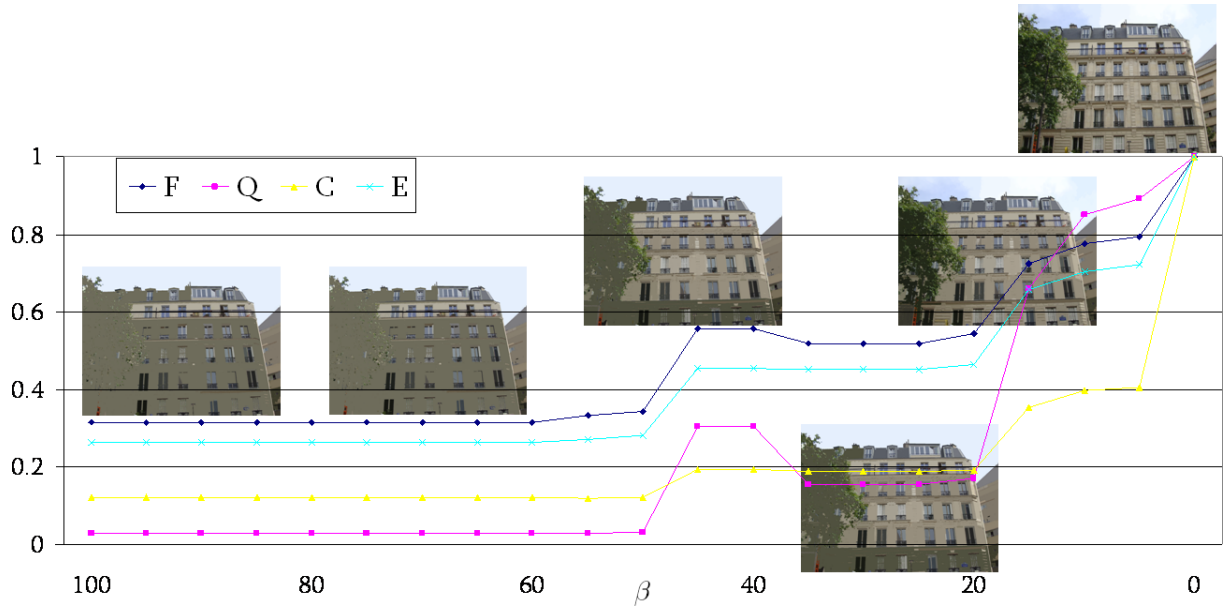


FIGURE 7.1 – Résultats des mesures de qualité d'une segmentation.

### Mesures normalisées d'évaluation

Comme nous l'avons mentionné précédemment, un des principaux problèmes des mesures de qualité d'une segmentation est le rapport de force entre le terme d'hétérogénéité couleur pour éviter une sous-segmentation et le terme de nombre de régions pour éviter une sur-segmentation. Si nous normalisons ces deux termes en utilisant l'information des deux pires cas de segmentation, nous pouvons avoir des indicateurs plus fiables. En analysant en détail les méthodes de segmentation basées sur le *Max-Tree*, nous connaissons les deux cas de segmentation extrema.

1. La pire sous-segmentation : la racine. Cette segmentation présentera le nombre de régions minimal ( $N_{\min} = 1$ ) et l'hétérogénéité couleur maximale ( $e_{\max}$ ).
2. La pire sur-segmentation : les feuilles. Cette segmentation présentera le nombre de régions maximal ( $N_{\max}$ ) et l'hétérogénéité couleur minimale ( $e_{\min}$ ).

Avec cette information nous établissons des mesures de segmentation normalisées ( $\bar{e}$  et  $\bar{n}$ ) comme l'illustre l'équation 7.1.

$$\begin{aligned}\bar{e} &= \frac{e - e_{\min}}{e_{\max} - e_{\min}} \\ \bar{n} &= \frac{N_R - 1}{N_{\max} - 1}\end{aligned}\tag{7.1}$$

où  $e$  est l'hétérogénéité couleur de l'image.

Nous proposons ainsi d'utiliser cette information afin d'estimer l'amélioration apportée par les approches proposées. La Figure 7.2 illustre l'évolution de ces mesures pour la même série de tests de la Figure 7.1. Nous pouvons observer le comportement suivant :

- Dans l'intervalle  $\beta = \{100, 60\}\%$  : le résultat de l'ouverture ultime hiérarchique est le même que l'ouverture ultime classique, car la région la plus grande occupe 60% de l'image.
- $\beta = \{55, 50\}\%$  : il y a une légère augmentation de  $\_n$ , mais  $\_e$  n'a pas tellement diminué.
- $\beta = \{45, 40\}\%$  : le nombre de régions a augmenté d'un facteur de 1,5 ( $\_n$ ), cependant  $\_e$  n'a diminué que d'un facteur de 0,7.
- $\beta = \{35, 20\}\%$  : le nombre de régions a diminué par rapport à l'intervalle précédent et  $\_e$  n'a pas tellement changé.
- $\beta = 15\%$  :  $\_e$  est réduit à un quart et  $\_n$  est augmenté d'un facteur de 2,5. Visuellement, le problème de masquage et des fuites a disparu.

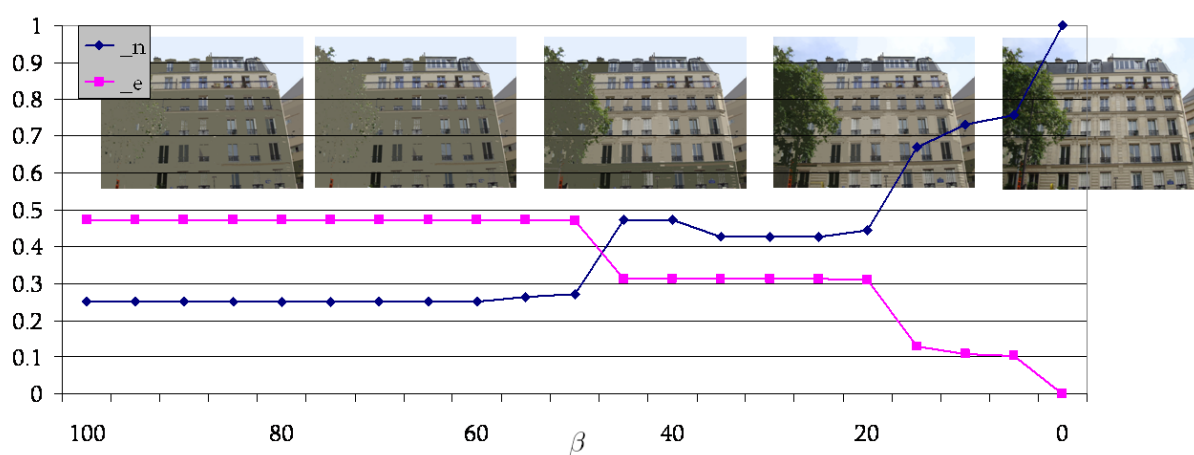


FIGURE 7.2 – Résultats des mesures normalisées de qualité d'une segmentation.

## 7.2.2 Ouverture ultime

Les méthodes ont été testées sur une base de 50 images urbaines. Néanmoins, nous illustrerons seulement les approches sur l'image de la Figure 7.3(a). Les autres résultats sont disponibles sur le site web suivant : [cmm.enscm.fr/~hernandez/UO/FacadeResults.html](http://cmm.enscm.fr/~hernandez/UO/FacadeResults.html).

Les résultats de l'opérateur classique  $\theta$  sont illustrés sur les Figures 7.3(c) et 7.3(d). Comme nous pouvons l'observer, l'opérateur masque plusieurs structures internes de la façade (des fenêtres, des volets, la porte) et fusionne la chaussée avec le bâtiment. L'indicatrice a 1953 régions et la différence couleur ( $e$ ) entre l'image mosaïque et

l'image originale est égale à 72. Les valeurs normalisées  $\_n$  et  $\_e$  sont respectivement 0,063 et 0,62.

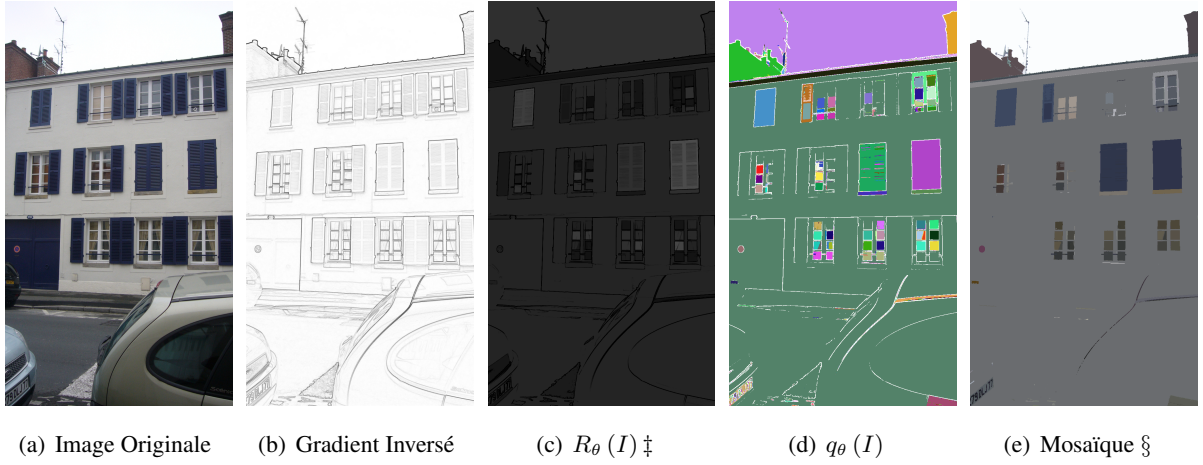


FIGURE 7.3 – (a) Image Originale, (b) Image Gradient Inversée, (c)-(d)  $R_\theta(I)$  et  $q_\theta(I)$  issues de l'ouverture ultime, (e) Image Mosaïque sur  $q_\theta(I)$ .  $\ddagger$  Correction gamma de 2.  $\S$  L'image mosaïque est la valeur moyenne de l'image originale sur la partition générée par la ligne de partage des eaux du gradient en utilisant comme marqueur  $q_\theta$ .

### 7.2.3 Ouverture ultime avec l'information de forme

Rappelons la fonction de forme pour les images de façade décrite par l'équation 7.2. Cette fonction a été désignée dans le but de favoriser les formes rectangulaires en utilisant l'attribut du facteur de remplissage. La fonction utilise également une relation hauteur-largeur et des contraintes de taille afin de ne pas favoriser les grandes structures.

$$f(\Omega) = 1 + \alpha \psi_{\kappa_1}(\Omega) \psi_{\kappa_2}(\Omega) \psi_{\kappa_3}(\Omega)$$

$$\psi_{\kappa_1}(\Omega) = (\Upsilon_\Omega)^2, \quad \psi_{\kappa_2}(\Omega) = \begin{cases} (\aleph_\Omega)^{0,5} & \aleph_\Omega > 0,2 \\ 0 & \text{sinon} \end{cases}, \quad \psi_{\kappa_3}(\Omega) = \begin{cases} 1 & h_\Omega < \frac{H_I}{3} \wedge w_\Omega < \frac{W_I}{3} \\ 0 & \text{sinon} \end{cases}$$

$$\Upsilon_\Omega = \frac{A_\Omega}{\text{Abbox}_\Omega}, \quad \aleph_\Omega = \frac{\min(h,w)}{\max(h,w)}, \quad H_I \text{ et } W_I \text{ hauteur et largeur de l'image} \quad \alpha = 9 \quad (7.2)$$

Les résultats de l'opérateur de forme  $\theta^\Omega$  sur l'image d'exemple sont présentés sur la Figure 7.4. Toutes les fenêtres-volets ont été bien détectées, ainsi que la porte. En comparant les images  $F_\theta^\Omega(I)$  et  $\widehat{F}_\theta^\Omega(I)$ , nous pouvons observer que d'autres structures de "bruit" ont un facteur différent de un ; mais elles continueront masquées grâce à leur faible contraste. Nous pouvons aussi apprécier, principalement dans les pixels des volets du rez-de-chaussée, que l'image de facteur maximal  $\widehat{F}_\theta^\Omega(I)$  est supérieure à l'image du facteur  $F_\theta^\Omega(I)$ . Ce comportement est dû au fait que



le facteur maximal n'arrive pas au même moment qu'un fort résidu. Il faut se rappeler que  $r^\Omega$  est la multiplication de ces deux termes, facteur et résidu. Comme prévu le bâtiment continue collé à la chaussée car la fonction n'a pas été conçue pour favoriser aucune de ces deux structures. Par ailleurs, le nombre de régions a augmenté à 2901 ( $\_n = 0,0938$ ) et la différence couleur a diminué à 59,19 ( $\_e = 0,49$ ).

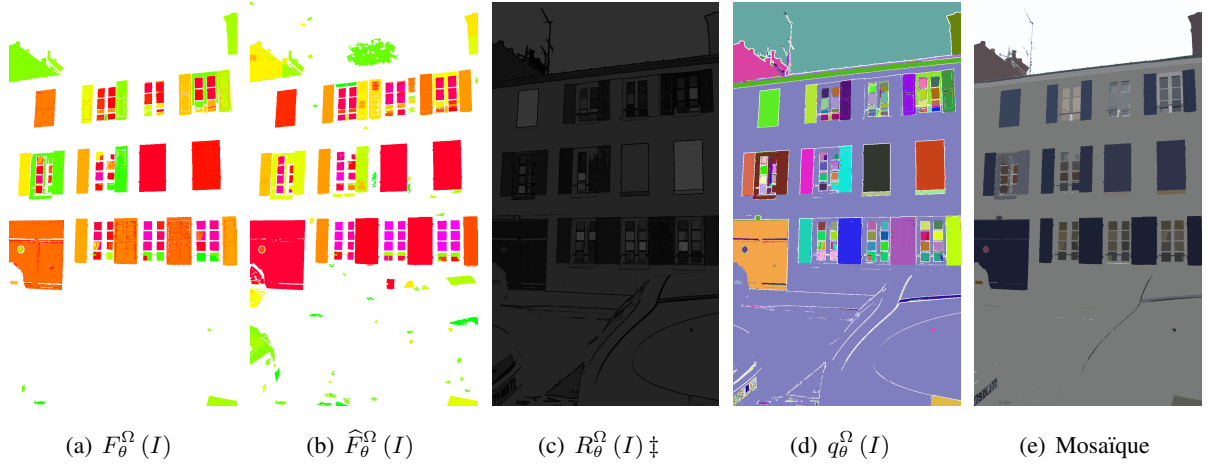


FIGURE 7.4 – (a)-(d)  $F_\theta^\Omega(I)$ ,  $R_\theta^\Omega(I)$  et  $q_\theta^\Omega(I)$  issues de l'ouverture ultime avec l'information de forme, (e) Image Mosaïque sur  $q_\theta^\Omega(I)$ .  
 $^\ddagger$  Correction gamma de 2.

## 7.2.4 Ouverture ultime hiérarchique

Nous avons déjà illustré dans la section 5.3 comment l'opérateur hiérarchique peut combattre les problèmes de l'ouverture ultime. Néanmoins, il nous reste à définir un critère hiérarchique pour une application générique. Nous avons déjà présenté deux critères, ainsi nous devons sélectionner plus précisément leur paramétrisation. Nous proposons de fixer les seuils de manière adaptative à partir des valeurs extraites de l'image (surface et écart-type) en utilisant un facteur multiplicatif  $\beta$ , ainsi :

- **Surface** :  $T = Surface(C_h^k) > \beta \times Surface(I)$
- **Hétérogénéité Couleur** :  $T = \Sigma \sigma_{\{R(C_h^k), V(C_h^k), B(C_h^k)\}} > \beta \times \Sigma \sigma_{\{R(I), V(I), B(I)\}}$

Sur notre base d'images de test, nous avons utilisé l'ouverture ultime hiérarchique en variant  $\beta$  depuis 100% jusqu'à 0% chaque 5%, pour chaque critère. Sur chaque image résultante, nous avons calculé les mesures d'évaluation  $\_e$  et  $\_n$ . Les Figures 7.5 et 7.6 présentent les moyennes des mesures  $\_e$  et  $\_n$ , pour les critères de surface et hétérogénéité couleur respectivement. La tendance est similaire pour les deux critères, car au fur et à mesure que le seuil diminue, l'hétérogénéité couleur diminue et le nombre de régions augmente. Après avoir regardé plusieurs images nous

avons constaté que les problèmes de masquage ont été éliminés avec une valeur de  $\beta = 15\%$  pour le critère de surface et une valeur  $\beta = 40\%$  pour le critère d'hétérogénéité couleur. Comme nous pouvons l'observer dans les Figures 7.5 et 7.6, avec ces valeurs nous obtenons des réductions moyennes de  $\_e$  d'un tiers par rapport à la version classique. D'autres valeurs de  $\beta$  pourront être choisies pour chaque application en fixant des contraintes de nombre de régions permis ou de différence couleur maximale acceptée.

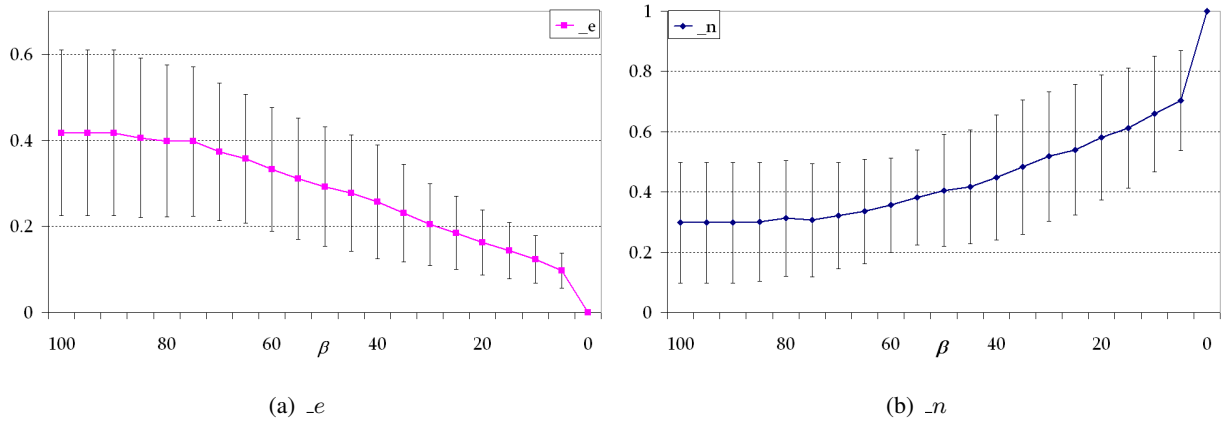


FIGURE 7.5 – Critère de surface. Mesures de qualité moyenne de l'ouverture ultime hiérarchique sur la base de test de 50 images.

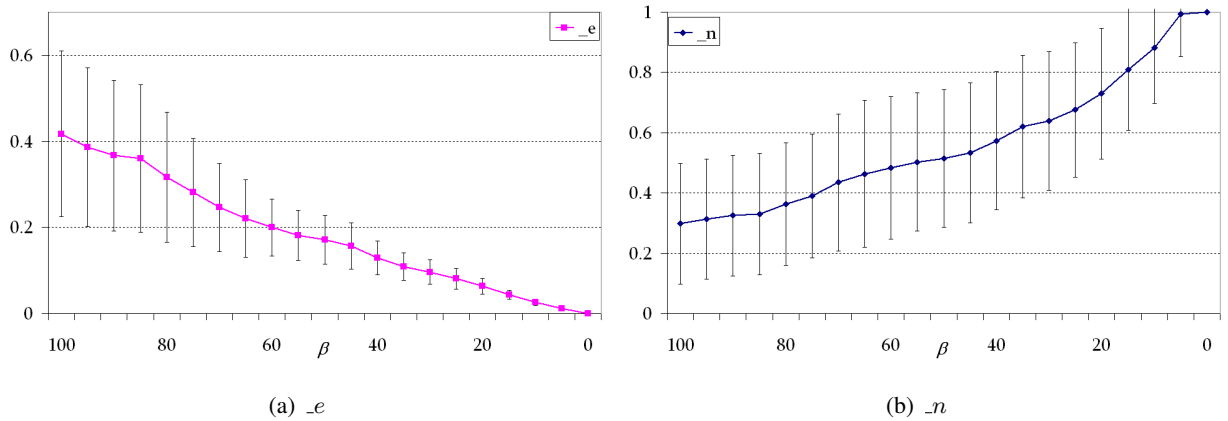


FIGURE 7.6 – Critère d'hétérogénéité couleur. Mesures de qualité moyenne de l'ouverture ultime hiérarchique sur la base de test de 50 images.

La Figure 7.7 illustre les résultats de l'opérateur hiérarchique sur l'image d'exemple. Nous pouvons observer que dans les deux cas, le problème de masquage et la connexion du bâtiment avec la chaussée ont disparu. De plus, les régions comme le ciel et le mur ne sont pas sur-segmentées. Visuellement, le critère d'hétérogénéité couleur présente de meilleurs résultats que le critère de surface : par exemple dans la partie basse de l'image les

voitures sont séparées de la chaussée. Pour cet exemple en particulier, avec les valeurs de  $\beta$  choisies, les mesures d'évaluation de la segmentation sont les suivantes :

	$\_n$	$\_e$
Surface	0, 231	0, 135
Hétérogénéité Couleur	0, 241	0, 098

Les valeurs montrent une meilleure mesure  $\_e$  pour le critère d'hétérogénéité couleur par rapport au critère de surface, avec presque le même nombre de régions, corroborant l'aspect visuel. Par rapport à l'ouverture classique la valeur  $\_e$  a diminué d'un facteur de 5 et  $\_n$  a augmenté d'un facteur de 4.

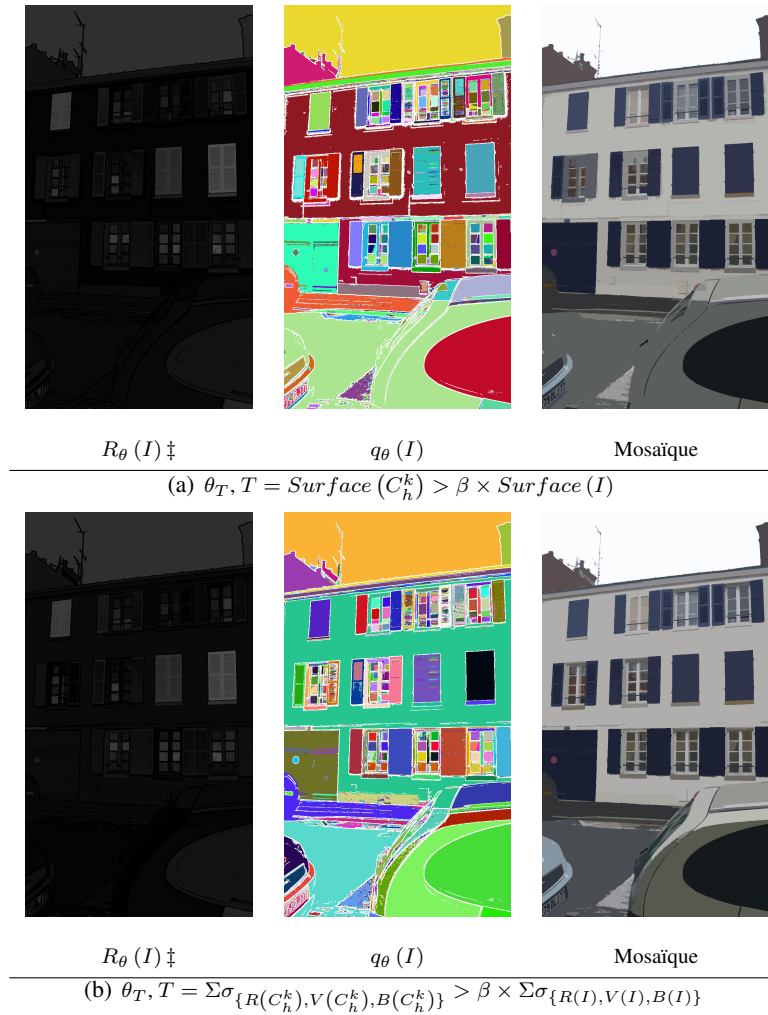


FIGURE 7.7 –  $R_\theta(I)$  et  $q_\theta(I)$  issues de l'ouverture ultime hiérarchique et image Mosaïque sur  $q_\theta$ .  $\ddagger$  Correction gamma de 2.

### 7.2.5 Ouverture ultime par $\Delta$ -attributs

Rappelons que nous appliquons l'opérateur sur une image gradient, afin d'appliquer l'ouverture ultime sur une polarité. Par ailleurs les structures de l'image originale présentent certaines transitions graduelles sur l'image gradient, c'est pour cela que l'approche avec accumulation est étudiée. D'abord, nous utilisons les mesures d'évaluation afin de déterminer le paramètre  $\Delta$  pour cette application. Sur la base de 50 images, nous avons utilisé l'opérateur en variant  $\Delta$  depuis 0 jusqu'à 20. La Figure 7.10 montre l'évolution moyenne de  $\_n$  et de  $\_e$ . Avec  $\Delta = 0$  l'opérateur devient l'ouverture ultime classique. De  $\Delta = 1$  à  $\Delta = 4$ , les valeurs de  $\_n$  et  $\_e$  diminuent légèrement et ensuite,  $\_n$  continue à décroître et  $\_e$  commence à augmenter. Le nombre de régions continue à diminuer car l'augmentation de  $\Delta$  produit la fusion de plus en plus de régions. Dans cet intervalle d'analyse la valeur moyenne de  $\_e$  est toujours inférieure à la valeur moyenne de l'ouverture ultime classique. Par contre, au-delà de  $\Delta = 20$ , la valeur moyenne de  $\_e$  augmente car les fusions de régions sont plus significatives. Pour toutes ces raisons nous choisissons une valeur de  $\Delta = 4$  pour l'application d'images de façade.

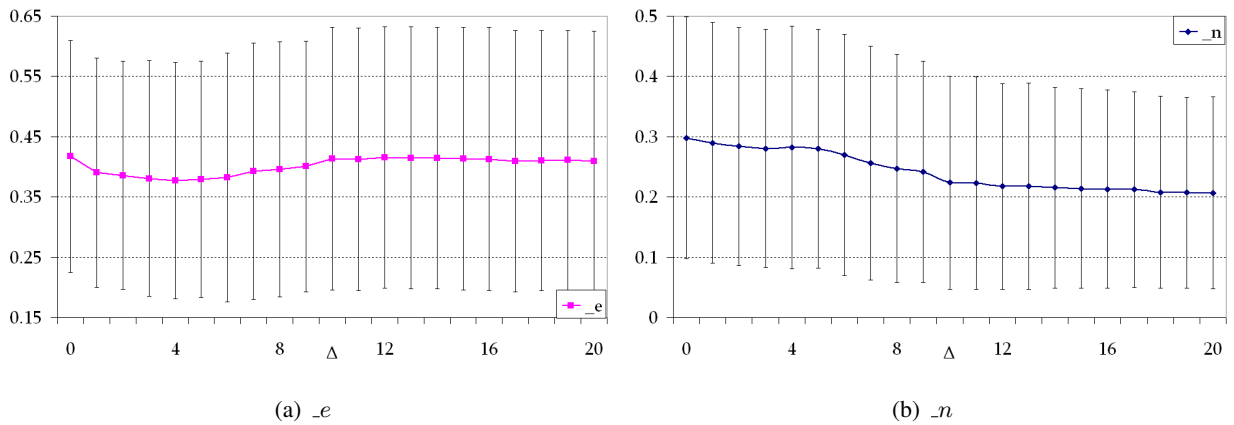


FIGURE 7.8 – Mesures de qualité moyenne de l'ouverture ultime par  $\Delta$ -attributs sur la base de test de 50 images.

Pour l'image de notre exemple, le résultat de l'ouverture ultime par  $\Delta$ -attribut, avec  $\Delta = 4$  est illustré sur la Figure 7.9. Nous pouvons observer plusieurs volets et vitraux supplémentaires détectés par rapport à l'opérateur classique. Cela explique la légère réduction dans la mesure  $\_e$ , car pour cet exemple  $\_e = 0,603$ . La valeur de  $\_n$  n'a pas changé ( $\_n = 0,063$ ). Nous pouvons aussi observer une augmentation considérable dans la transformée ; par exemple la région du ciel est passée de 47 à 70 et des fenêtres sont passées de 50 à 90 approximativement. Ce qui confirme le principal avantage de l'opérateur avec accumulation, une meilleure estimation du contraste malgré les transitions graduelles.

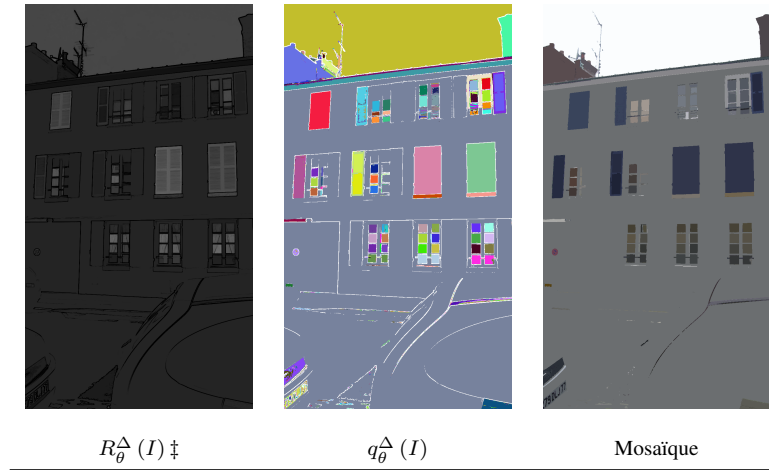


FIGURE 7.9 –  $R_{\theta}^{\Delta}(I)$  et  $q_{\theta}^{\Delta}(I)$  issues de l'ouverture ultime hiérarchique et image Mosaïque sur  $q_{\theta}^{\Delta}$ . ‡ Correction gamma de 2.

### 7.2.6 Discussion

Les mesures moyennes pour les trois approches sont illustrées dans la Figure 7.10. Les trois approches présentées réduisent les problèmes de l'ouverture ultime ; c'est pourquoi nous pouvons observer une réduction de la valeur moyenne  $\_e$  pour les trois cas. Cependant cette diminution implique une augmentation du nombre de régions. La seule approche qui réduit en même temps le nombre de régions et l'hétérogénéité couleur est l'opérateur avec accumulation. Les plus grandes réductions de  $\_e$  sont présentées dans les approches hiérarchiques mais aussi accompagnées du plus grand nombre de régions. Par ailleurs il ne faut pas oublier l'information additionnelle que produit l'opérateur avec la forme. Cette information pourrait être facilement utilisée dans une procédure de reconnaissance.

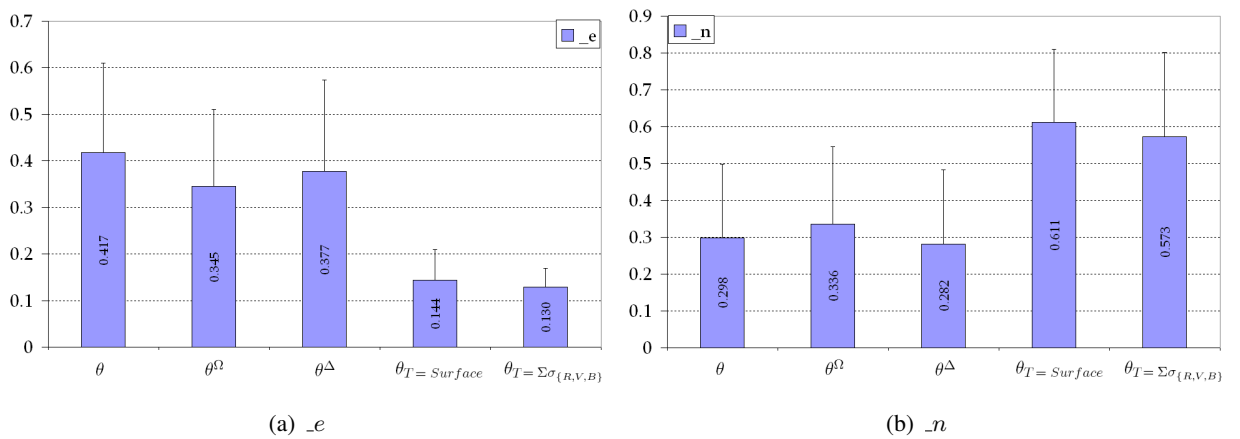


FIGURE 7.10 – Mesures de qualité moyenne sur la base de test de 50 images.

### 7.3 Détection de texte enfoui

Nous avons déjà présenté l'utilisation de l'ouverture ultime pour la segmentation de texte enfoui dans la section 6.3. Nous avons calculé l'opérateur sur la composante de luminance (une polarité) de l'image couleur ; cependant, afin de détecter le texte enfoui sur des images, les deux polarités de l'image doivent être analysées. Un exemple d'une image avec du texte clair sur un fond sombre et vice-versa est illustré sur la Figure 7.11.



FIGURE 7.11 – Exemple d'image de texte dans les deux polarités.

Nous proposons d'utiliser l'ouverture ultime sur chaque polarité séparément et d'ajouter une dernière étape de sélection des pixels basée sur la transformée. Pour chaque pixel la polarité donnant lieu à la plus forte transformée est sélectionnée (voir équation. 7.3).

$$\begin{aligned}
 R_{\theta}(I, I^c) &= \max(R_{\theta}(I), R_{\theta}(I^c)) \\
 q_{\theta}(I, I^c) &= \begin{cases} q_{\theta}(I) & R_{\theta}(I) > R_{\theta}(I^c) \\ q_{\theta}(I^c) & \text{sinon} \end{cases}
 \end{aligned} \tag{7.3}$$

Pour la détection de texte enfoui, nous avons additionné trois «filtres» afin de garder l'information la plus pertinente dans la sortie de l'opérateur. Le premier filtrage supprime les composantes mal évaluées de l'indicateur (valeurs perchées). Ensuite nous éliminons les petites régions, en utilisant un seuillage par surface de 10 pixels. Finalement, les régions de faible contraste sont supprimées. Nous calculons la transformée moyenne sur chaque région de l'indicateur et nous conservons les régions dont le contraste est supérieur à 5 niveaux de gris. Le résultat de l'ouverture ultime sur chaque polarité en appliquant les trois filtres proposés est présenté sur la Figure 7.12.

Cette section de segmentation de texte enfoui se déroule de la manière suivante. D'abord nous allons présenter la sélection des paramètres pour chacune des approches proposées et ensuite nous ferons une comparaison visuelle de la détection de texte enfoui. Finalement, nous discuterons de l'efficacité des méthodes.

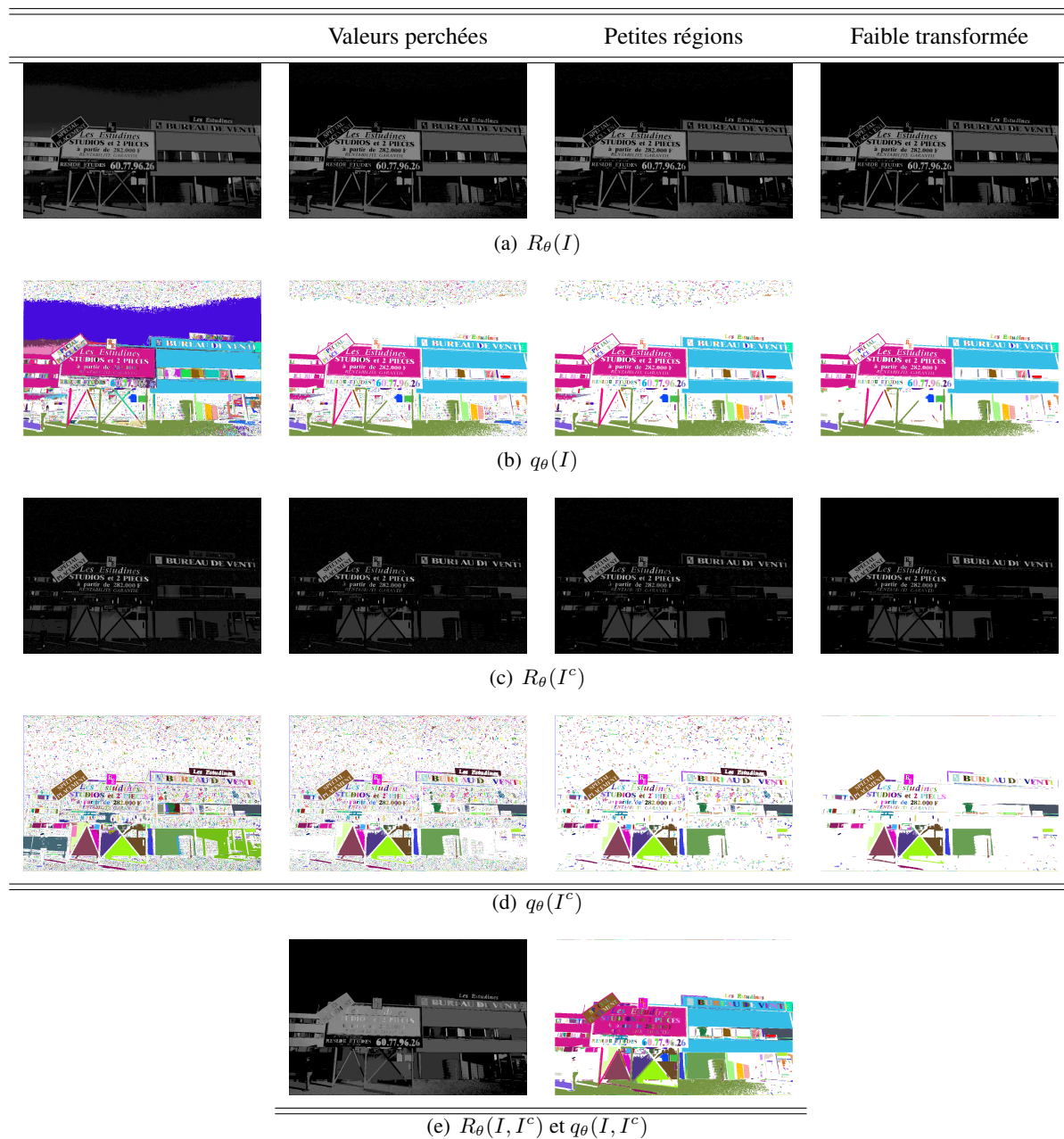


FIGURE 7.12 – Exemple de l'ouverture ultime pour la détection de texte enfoui sur les deux polarités de l'image.

### 7.3.1 Ouverture ultime avec l'information de forme

Le premier pas avant d'utiliser l'opérateur avec la forme est d'établir les caractéristiques *a priori* pour des formes d'intérêt. Nous avons analysé environ 5000 caractères et nous avons extrait plusieurs attributs : hauteur, largeur, surface, etc... De la même manière que nous l'avons fait pour l'application de façade, nous proposons d'utiliser les attributs du facteur remplissage et le rapport hauteur-largeur pour la définition de la fonction de forme, car ces attributs sont invariants à l'échelle. La Figure 7.13 illustre les histogrammes de ces deux attributs dont nous pouvons déduire l'information suivante :

- 97% de lettres ont un rapport hauteur-largeur ( $\aleph$ ) plus grand que 0,4.
- 84% de lettres ont un facteur de remplissage ( $\Upsilon$ ) entre 0,2 et 0,9. (centre  $\Upsilon_{\Omega_{ref}} = 0,55$  et limites  $\tau_{\Upsilon} = \pm 0,35$  du centre.)

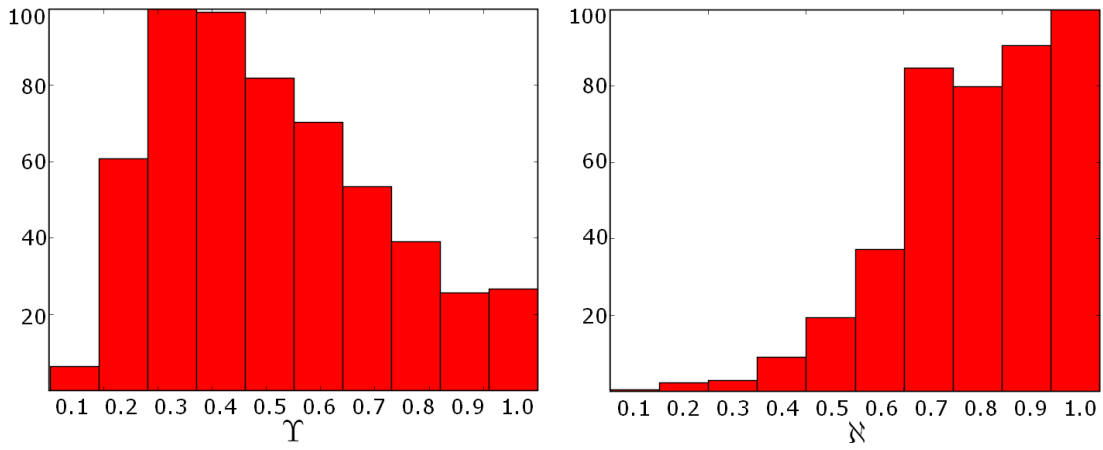


FIGURE 7.13 – Histogrammes de 5000 caractères analysés.

Avec cette information, nous avons utilisé les fonctions de similarité de forme de l'équation 4.1, page 49, avec une réponse de la fonction de large réponse ( $\varsigma_{\kappa} = 0,5$ ). Nous avons aussi utilisé deux contraintes de taille, car la plus grande hauteur et largeur d'un caractère est un tiers de la hauteur ( $H_I$ ) et de la largeur ( $W_I$ ) de l'image, respectivement. L'équation 7.4 présente la fonction du facteur de forme que nous avons utilisé pour l'application de détection de texte.

$$f(\Omega) = 1 + \alpha \psi_{\kappa_1}(\Omega) \psi_{\kappa_2}(\Omega) \psi_{\kappa_3}(\Omega)$$

$$\psi_{\kappa_1}(\Omega) = \begin{cases} (\aleph_{\Omega})^{0,5} & \aleph_{\Omega} > 0,4 \\ 0 & \text{sinon} \end{cases}, \quad \psi_{\kappa_2}(\Omega) = \begin{cases} \left(1 - \left(\frac{\Upsilon_{\Omega} - 0,55}{0,35}\right)^{0,5}\right) & 0,2 < \Upsilon_{\Omega} < 0,9 \\ 0 & \text{sinon} \end{cases}, \quad (7.4)$$

$$\psi_{\kappa_3}(\Omega) = \begin{cases} 1 & h_{\Omega} < \frac{H_I}{3} \wedge w_{\Omega} < \frac{W_I}{3} \\ 0 & \text{sinon} \end{cases} \quad \alpha = 9$$



### 7.3.2 Ouverture ultime hiérarchique

Dans cette application, nous proposons aussi d'utiliser pour l'approche hiérarchique avec les deux critères, surface et hétérogénéité de couleur.

- **Surface** :  $T = Surface(C_h^k) > \beta \times Surface(I)$
- **Hétérogénéité Couleur** :  $T = \Sigma \sigma_{\{R(C_h^k), V(C_h^k), B(C_h^k)\}} > \beta \times \Sigma \sigma_{\{R(I), V(I), B(I)\}}$

Nous avons fixé le seuil de surface à partir de l'analyse des histogrammes, car la plupart des caractères ont une surface inférieure à un dixième de la surface de l'image ( $\beta = 10\%$ ). Concernant le seuil de l'hétérogénéité couleur, nous avons fait l'hypothèse que les caractères ont une couleur homogène, donc nous avons fixé le seuil avec une valeur de  $\beta = 30\%$ .

### 7.3.3 Ouverture ultime par $\Delta$ -attributs

Pour l'utilisation de l'ouverture ultime par  $\Delta$ -attributs, nous ne considérons que les valeurs de  $\Delta = 1$  et  $\Delta = 2$ , afin d'éviter la fusion de caractères, car en plusieurs cas les caractères sont trop proches les uns des autres. Ainsi, avec une valeur de  $\Delta$  supérieure à 2 la fusion entre caractères est plus probable.

### 7.3.4 Résultats

Afin de tester les différentes méthodes, une base d'approximativement 500 d'images a été utilisée. Nous allons présenter les résultats des trois approches sur trois images ; cependant, le lecteur peut se référer au site web [cmm.ensmp.fr/~hernandez/UO/TextResults.html](http://cmm.ensmp.fr/~hernandez/UO/TextResults.html) pour observer les autres résultats. La Figure 7.14 illustre les trois images originales choisies.

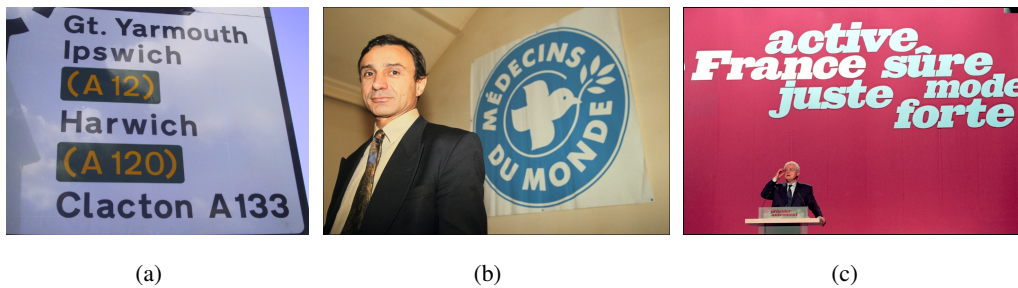


FIGURE 7.14 – Images de test pour la détection de texte enfoui.

Les résultats de la Figure 7.14(a) utilisant les ouvertures ultimes sont présentés sur la Figure 7.15, dont nous pouvons observer le comportement suivant :

$\theta$  : La plupart des lettres ont été détectées, sauf les caractères de la polarité claire sur un fond sombre : “(A12)» et “(A120)», car le contraste entre le rectangle sombre et le panneau blanc est plus fort que le contraste entre le rectangle sombre et les lettres jaunes à l’intérieur. Dans les cas particulier de “(A12)», le “2» et le “)”» sont détectés en partie car ils ont une partie plus éclairée. Cette variation de la luminance fractionne aussi la lettre “m» de “Yarmounth».

$\theta^\Omega$  : L’ouverture ultime avec information de forme corrige le problème de masquage, sauf pour le chiffre “1»». Ce chiffre n’est pas favorisé par la fonction de la forme car son attribut  $\aleph$  est inférieur à 0,4. Au niveau de la transformée, les lettres sont les structures détectées les plus contrastées. D’ailleurs ce contraste est supérieur à celui obtenu par l’opérateur classique. Cependant des “i» et des “l» n’ont pas été favorisés pour la même raison que le chiffre “1»».

$\theta_T$  : L’utilisation de la hiérarchie avec le critère de surface ne résout pas le masquage des lettres jaunes, car les rectangles sombres ont une surface inférieure à un dixième de la surface de l’image. Par contre, avec le critère d’hétérogénéité couleur, la plupart des lettres sont détectées. Par ailleurs, avec ce critère, les lettres qui ont des variations de luminance à cause des ombres sont coupées, par exemple la lettre “m» de “Yarmounth» et la lettre “w» de “Ipswich».

$\theta^\Delta$  : Avec l’ouverture ultime par  $\Delta$ —attribut le problème de masquage continue et seulement quelques pixels des lettres jaunes deviennent visibles. Par ailleurs, la lettre “m» de “Yarmounth» n’est plus cassée. Le contraste des structures détectées est plus fort que celui présenté par l’opérateur classique.

La Figure 7.16 illustre les résultats des ouvertures ultimes de la Figure 7.14(b).

$\theta$  : L’ouverture ultime sur cette image produit un fort masquage de toutes les structures d’intérêt.

$\theta^\Omega$  : L’ouverture ultime avec l’information de forme détecte toutes les lettres. Sauf la lettre “1» qui continue d’être masquée. Le facteur de forme favorise également d’autres structures comme les yeux et les cheveux de l’individu.

$\theta_T$  : Comme prévu l’approche hiérarchique résout le problème de masquage, car dans cet exemple le problème est produit par une région (nœud du *Max-Tree*) très hétérogène et qui a une grande surface. En comparant l’ouverture hiérarchique ( $\theta_T$ ) avec l’information forme ( $\theta^\Omega$ ), l’oiseau est détecté. Dans le cas de  $\theta^\Omega$  l’oiseau n’a pas un bon facteur de forme et il ne sera pas détecté, alors que dans les cas de  $\theta_T$ , l’oiseau est au même niveau hiérarchique que les lettres.

$\theta^\Delta$  : Seules certains pixels de lettres deviennent visibles. Le nombre de pixels visibles est plus important avec une valeur de  $\Delta = 2$ .

Finalement la Figure 7.17 illustre les résultats des ouvertures ultimes de la Figure 7.14(c).

$\theta$  : L’ouverture ultime détecte les plus grosses lettres ; néanmoins, les lettres “ranc» de “France» sont supprimées par le filtre de valeurs perchées. En analysant l’ouverture ultime comme une procédure séquentielle, les lettres

“rance» ont été initialement segmentées dans une seule structure. Ensuite, la lettre “e» est mise à jour avec un résidu plus fort que la valeur courante de la transformée. Ainsi, les lettres “ranc» ont une mauvaise évaluation car leur indicatrice correspond à celle des lettres “rance». En outre, l’opérateur masque des caractères sur la tribune de l’orateur.

$\theta^\Omega$  : L’ouverture ultime avec information de forme résout le problème de mauvaise évaluation de “ranc», en segmentant les lettres séparément. C’est un bon exemple du problème de fuites dans le texte dont l’opérateur avec la forme améliore le résultat. L’opérateur détecte également les lettres dans la tribune.

$\theta_T$  : L’ouverture ultime hiérarchique détecte les lettres dans la tribune. Néanmoins, le problème de mauvaise évaluation continue et les lettres “ranc» sont supprimées par le filtre des valeurs perchées.

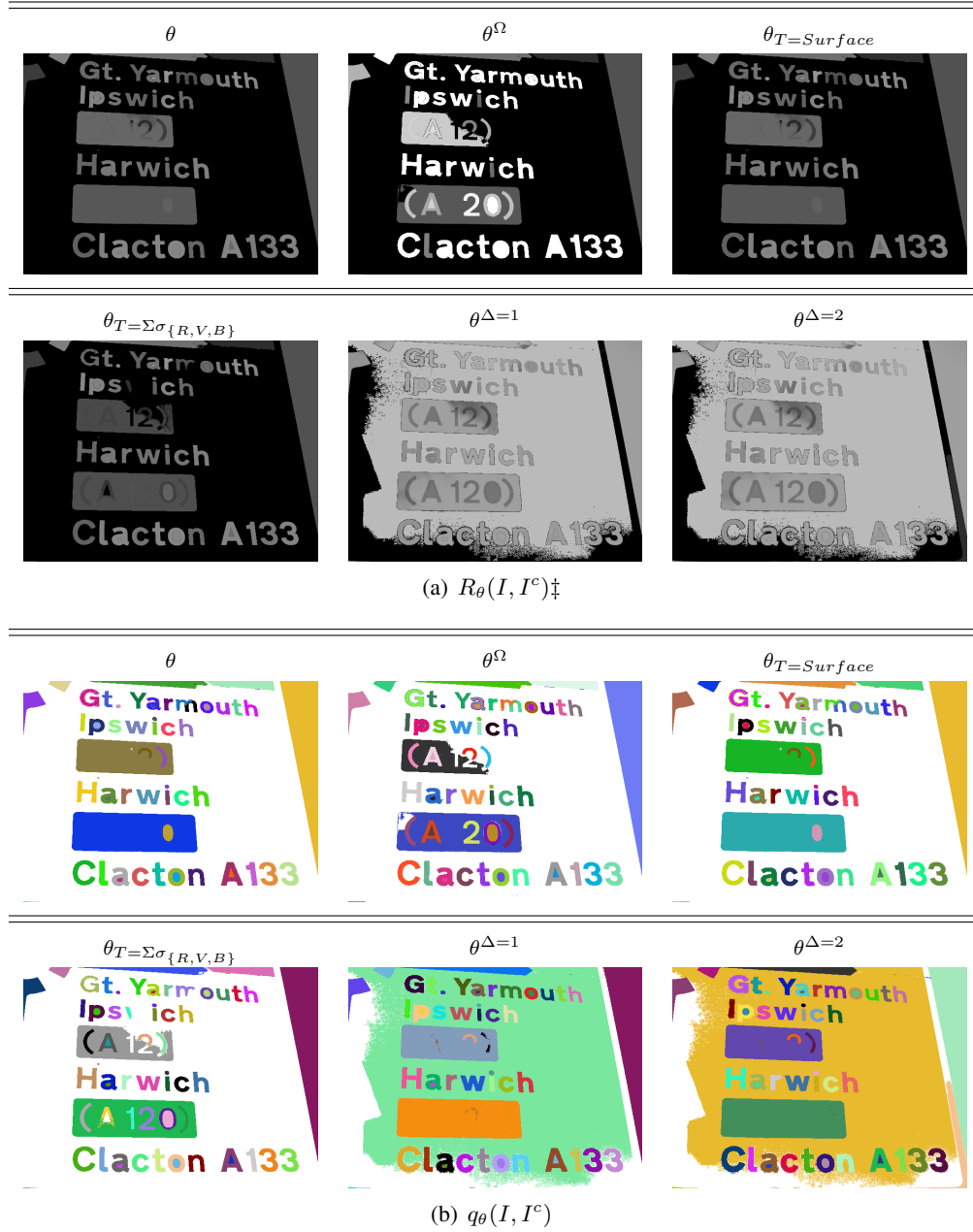
$\theta^\Delta$  : Dans cette image le fond est en dégradé. En accumulant les résidus différents de zéro, le résidu du fond devient plus fort que le résidu de certaines lettres. Ainsi, ces lettres sont masquées.

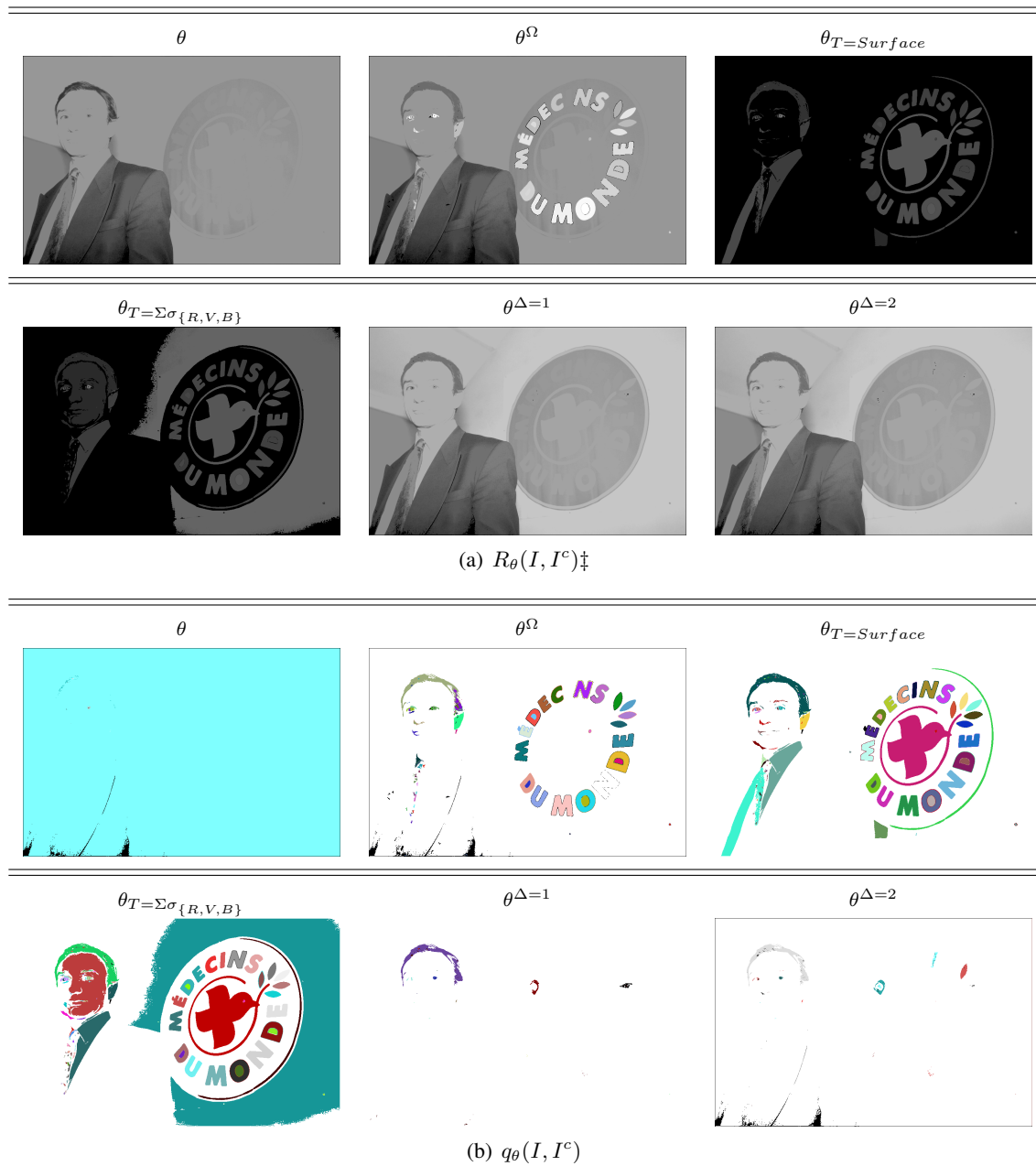
### 7.3.5 Discussion

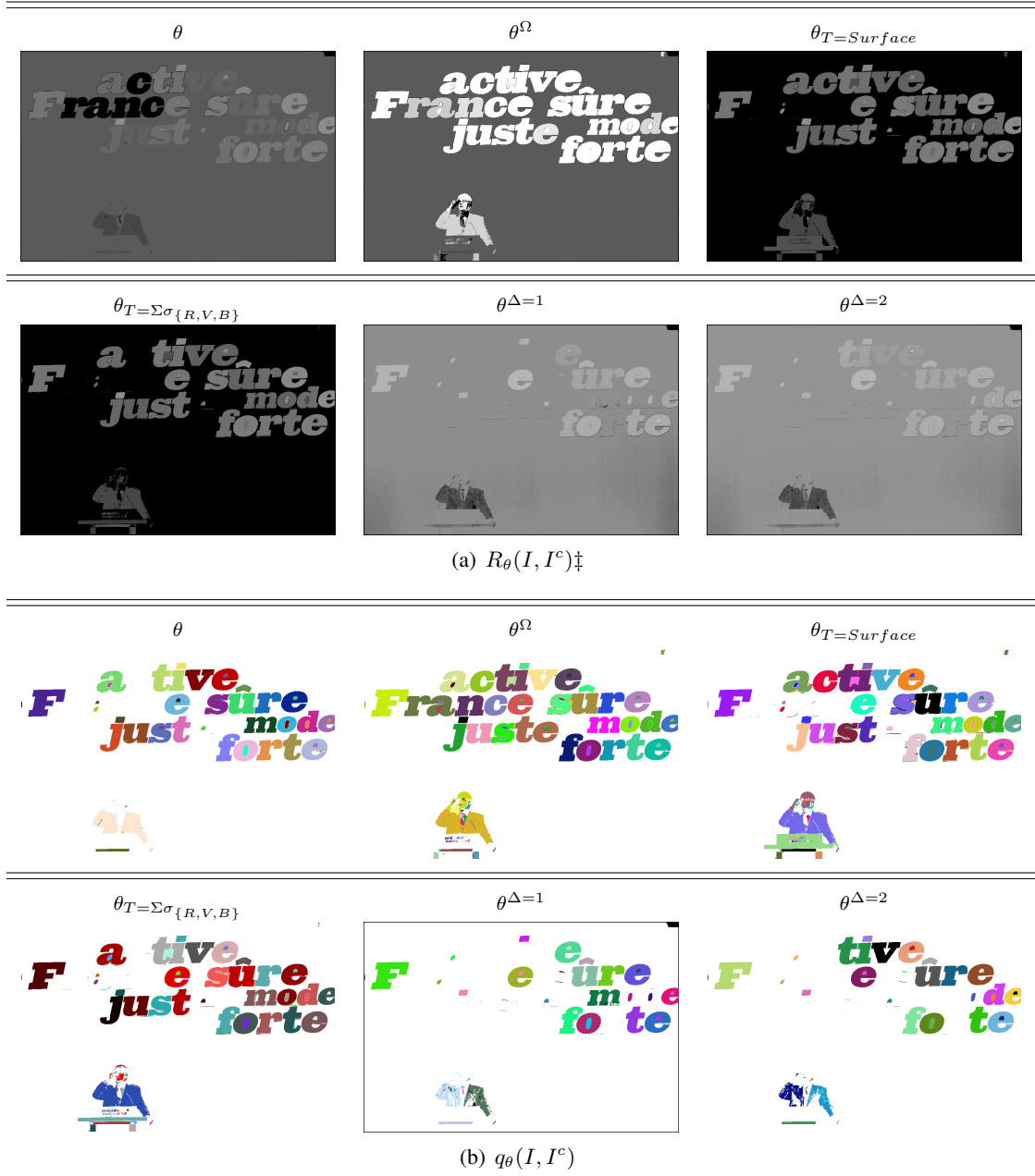
Nous venons de présenter l’utilisation de l’ouverture ultime pour une application de détection de texte enfoui. Nous avons illustré les différents cas problématiques de l’opérateur et les améliorations apportées lorsque nous utilisons les méthodes proposées.

Concernant l’ouverture ultime avec l’information de forme, elle peut réduire à la fois le problème de masquage et des fuites. Nous avons paramétré la fonction de forme à l’aide d’une analyse de plusieurs caractères. Cette fonction favorise la majorité des caractères. Cependant des caractères allongés et complètement remplis, tels que le “i» et “l», ne sont pas pris en compte par notre fonction. Nous proposons d’ajouter à cette fonction un terme pour favoriser ce type de caractères. Par ailleurs, les améliorations de l’opérateur sont visibles au niveau de l’indicatrice et de la transformée, car il assigne la bonne évaluation avec un fort contraste.

A propos de l’approche hiérarchique, elle combat principalement le problème de masquage. Nous avons présenté un exemple où le problème de masquage n’a pas été résolu car les lettres étaient contenues à l’intérieur d’une structure qui ne vérifiait pas le critère hiérarchique de surface (voir Figure 7.17). En outre, nous avons aussi proposé un critère d’hétérogénéité couleur basé sur les composantes *RVB*. Cependant ce critère est sensible aux variations de luminance dans l’image, c’est pourquoi nous proposons d’étudier d’autres critères d’hétérogénéité couleur basés sur les composantes de chrominance. Par ailleurs, les effets de cette amélioration sont visibles au niveau de l’indicatrice et de la transformée, car les structures qui produisent le masquage ont été re-segmentées. D’ailleurs au niveau de la transformée le contraste n’est pas avantageé comme dans les autres approches. C’est pour cela que nous pouvons imaginer une combinaison de cette approche avec l’accumulation de résidus, pour donner aux caractères un contraste plus important.


 FIGURE 7.15 – Résultats des ouvertures ultimes sur la Figure 7.14(a).  $\ddagger$  Correction gamma de 2.


 FIGURE 7.16 – Résultats des ouvertures ultimes sur la Figure 7.14(b).  $\ddagger$  Correction gamma de 2.


 FIGURE 7.17 – Résultats des ouvertures ultimes sur la Figure 7.14(c).  $\ddagger$  Correction gamma de 2.

## 7.4 Segmentation d'images de cellules

### 7.4.1 Introduction

L'émergence récente de technologies d'acquisition automatisée d'images de haut débit a changé la façon dont les biologistes cellulaires recueillent et analysent les données [176]. L'interprétation des phénotypes cellulaires est maintenant un important outil d'investigation en biologie. Cette étude est réalisée dans le cadre du projet RAMIS Cancer-Bio-Santé Pôle de Toulouse-France. L'objectif du projet est de développer une stratégie innovante basée sur l'imagerie haute résolution pour la caractérisation de protéines impliquées dans la division de cellules cancéreuses humaines et de molécules modulatrices. Par marquage multi-paramétrique des cellules et microscopie automatisée seront générées des collections d'images représentatives de phénotypes induits par différents traitements pharmacologiques connus (molécules, RNAi). Un outil d'analyse automatique permettra de générer un profil phénotypique associé à ces différents traitements, reflétant ainsi leur mode d'action.

En traitement d'images, les propriétés morphologiques (taille, forme, texture, etc) de la cellule fournissent des informations précieuses qu'on peut explorer à partir d'une étape de segmentation. La proximité et le flou de cellules donnent lieu à un problème de fuites connu de l'ouverture ultime. L'utilisation de l'approche de forme et l'approche avec accumulation permettront de pallier ces problèmes. Concernant l'approche avec accumulation nous n'utiliserons qu'une valeur de  $\Delta = 1$  pour ne pas accentuer le problème de fuites. L'approche hiérarchique ne sera pas utilisée car le problème dû aux structures imbriquées n'a pas été identifié dans la segmentation d'images de cellules.

### 7.4.2 Ouverture ultime avec l'information de forme

Afin de déterminer la fonction du facteur, nous avons adopté la même technique utilisée par l'application de texte. Nous avons analysé environ 1200 cellules dont nous pouvons établir la circularité comme l'attribut le plus représentatif. Par ailleurs, les images sur lesquelles nous travaillons sont prises avec les mêmes conditions et les cellules se trouvent à peu près à la même échelle. Ainsi, nous avons composé la fonction de forme avec des contraintes de taille : surface, largeur et hauteur. L'équation 7.5 définit la fonction du facteur de forme que nous avons choisi.

$$\begin{aligned}
f(\Omega) &= 1 + \alpha \psi_{\kappa_1}(\Omega) \psi_{\kappa_2}(\Omega) \psi_{\kappa_3}(\Omega) \\
\alpha &= 9 \qquad \psi_{\kappa_1}(\Omega) = \frac{4\pi A_\Omega}{(L_\Omega)^2} \\
\psi_{\kappa_2}(\Omega) &= \begin{cases} 1 & 500 < A_\Omega < 3000 \\ 0 & \text{sinon} \end{cases} \quad \psi_{\kappa_3}(\Omega) = \begin{cases} 1 & 15 < h_\Omega < 110 \\ & \text{and} \\ & 15 < w_\Omega < 110 \\ 0 & \text{sinon} \end{cases} \quad (7.5)
\end{aligned}$$

### 7.4.3 Résultats

Les méthodes ont été testées sur une base de 20 images de cellules. Nous avons aussi utilisé un filtrage des composantes petites (élimination de régions de surface inférieure à 10 pixels) et de faible transformée (élimination de régions de transformée moyenne inférieure à 2 niveaux de gris). Les résultats sur une image sont illustrés sur la Figure 7.18. Les autres résultats sont disponibles sur le site web suivant : [cmm.enscm.fr/~hernandez/UO/CellResults.html](http://cmm.enscm.fr/~hernandez/UO/CellResults.html).

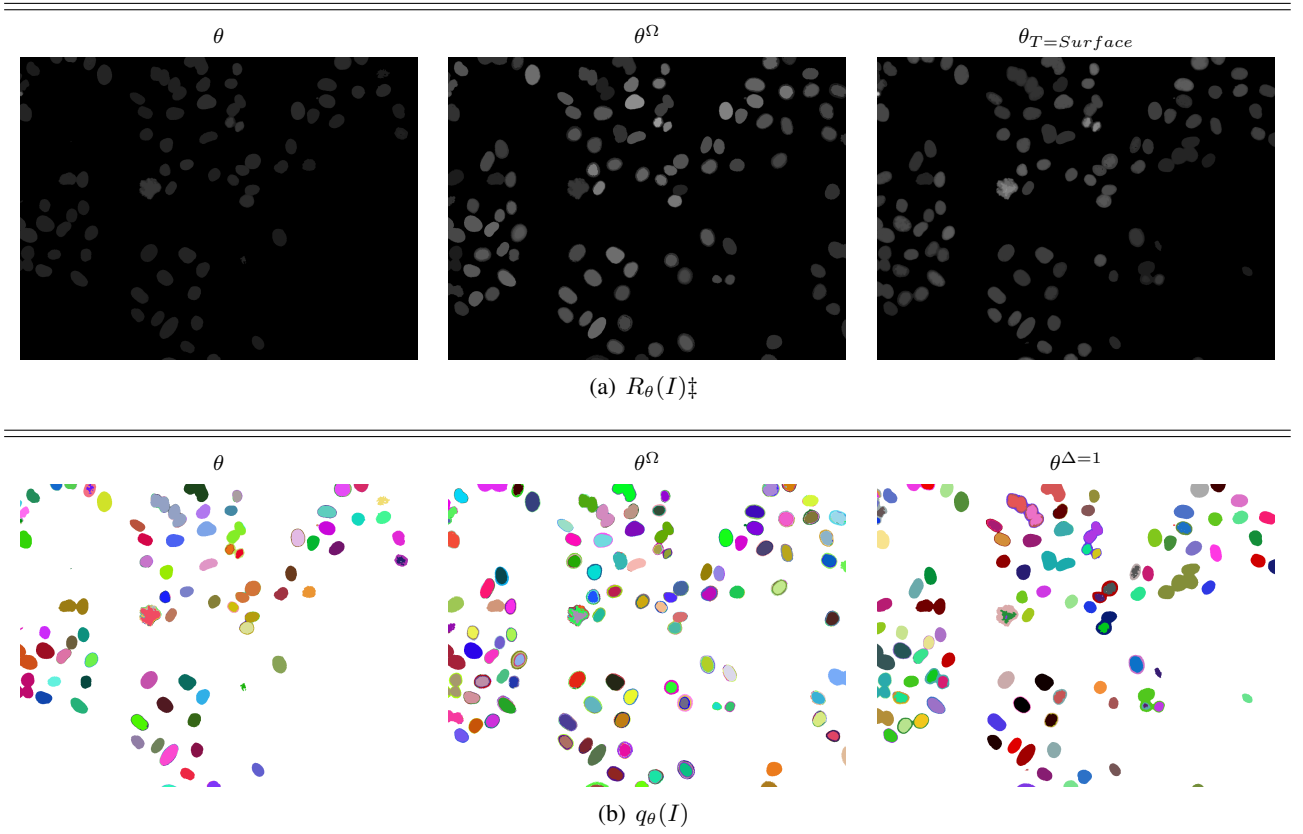
En comparant les indicatrices des trois méthodes, nous pouvons apprécier que les deux approches proposées détectent plus de cellules par rapport à la version classique. Au niveau de la transformée aussi nos approches produisent un contraste plus représentatif. Dans le cas particulier de la version avec la forme plusieurs cellules, qui étaient fusionnées, ont été séparées. De plus la sur-segmentation du noyau a été éliminée. Par ailleurs, une espèce d'anneaux, liés au flou des cellules, apparaît sur certaines cellules, car ces anneaux ont un fort facteur de forme à cause de leur circularité. Concernant l'approche avec accumulation, elle réduit aussi la sur-segmentation du noyau ; cependant, elle crée des connexions entre cellules trop proches.

### 7.4.4 Discussion

Nous avons présenté l'utilisation de l'ouverture ultime avec l'information de forme et l'ouverture ultime par  $\Delta$ -attribut pour la segmentation d'images de cellules. Les deux approches présentent de meilleurs résultats par rapport à l'opérateur classique car elles détectent plus de cellules et présentent un plus fort contraste au niveau de la transformée. Même si les résultats sont satisfaisants, nous proposons d'étudier la combinaison de deux approches, car plusieurs cellules ont un fort contraste et un faible facteur et vice-versa.

Par ailleurs, l'approche avec l'information de forme peut éviter la fusion de cellules lorsqu'un fort facteur arrive avant la fusion. La Figure 7.19 présente une série de seuils où nous observons les composantes de cellules au fur et à mesure que le seuil descend. Avant la fusion, aucune des deux cellules n'a un facteur suffisamment fort pour



FIGURE 7.18 – Résultats des ouvertures ultimes sur une image de cellules.  $\ddagger$  Correction gamma de 2.

segmenter les cellules de manière séparée.

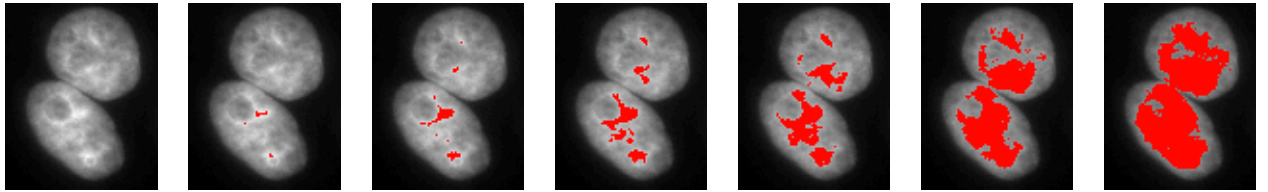


FIGURE 7.19 – Evolution de différents seuils de manière décroissante.

## 7.5 Discussion Générale

Dans ce chapitre nous avons présenté l'utilisation de l'ouverture ultime et trois améliorations proposées dans le cadre de trois applications : segmentation d'images de façade, détection de texte enfoui et segmentation de cellules. Jusqu'à présent nous avons présenté des résultats de segmentation. Ainsi, une étape de reconnaissance devrait être

rajoutée pour compléter la chaîne de traitement.

Parmi les trois méthodes présentées, l'opérateur avec l'information de forme est celui qui présente les résultats les plus intéressants vis-à-vis de la segmentation car cette approche réduit le problème de masquage ainsi que le problème de fuites, sans augmenter tellement le nombre de régions. De plus, l'opérateur de forme fournit une image additionnelle. D'ailleurs, même si les méthodes donnent de résultats intéressants, la combinaison de ceux-ci est une voie d'étude prometteuse qui pourrait être abordée par la suite.

### 7.5.1 Image d'entrée

Jusqu'à présent, nous avons utilisé les opérateurs (classique et les extensions) sans faire tellement attention à l'image d'entrée utilisée. L'ouverture ultime, comme son nom l'explique, extrait les composantes **claires** les plus contrastées sur un fond sombre. C'est pourquoi l'application de l'opérateur sur l'image inversée est aussi recommandée afin de trouver les composantes **sombres** sur un fond clair.

Par ailleurs, le fait d'utiliser l'image gradient inversée évite l'analyse de ces deux polarités, en réduisant le temps de calcul. Cependant l'utilisation de l'image gradient est réservée aux images avec des structures des contours plutôt nets. Voyons l'application de l'opérateur sur une image floue et sur son gradient inversé (voir Figure 7.20). Sur l'image de luminance, l'ouverture ultime détecte la lettre en diverses structures, alors que sur l'image gradient inversée il y a principalement de bruit. Ce comportement est produit car le gradient de la transition produit plusieurs passages à zéro (255 pour le gradient inversé, voir vue 3D du gradient), partitionnant la lettre.

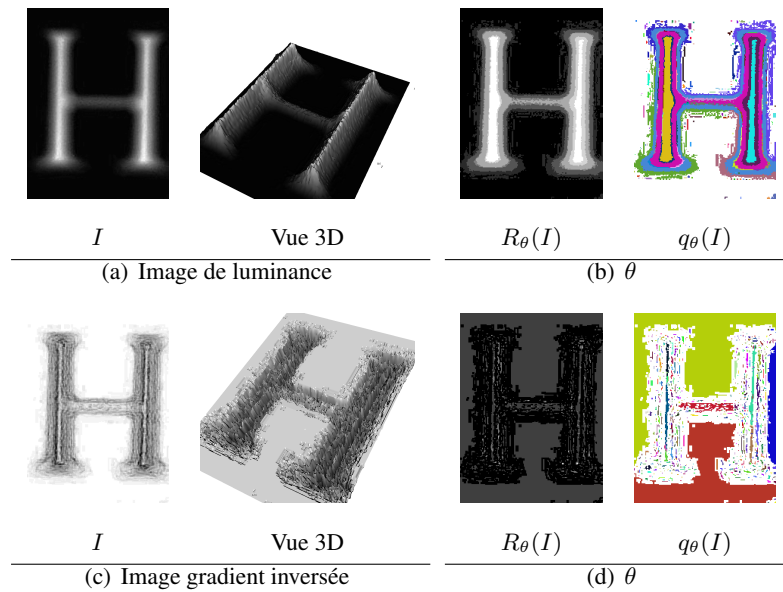


FIGURE 7.20 – Analyse de l'ouverture ultime sur l'image de luminance et l'image gradient.

En conclusion, l'utilisation de l'ouverture ultime sur l'image gradient inversée est vivement déconseillée lorsque l'image a des structures floues ou très fines. C'est pourquoi, dans l'application des images de texte enfoui, nous avons utilisé l'image de luminance sur chaque polarité, en choisissant pour chaque pixel la polarité donnant lieu au plus fort contraste. Concernant l'application de façades nous travaillons sur l'image de gradient. Finalement, sur l'application de cellules nous appliquons l'ouverture ultime sur une seule polarité car les cellules sont toujours plus claires que le fond.

### 7.5.2 Paramétrage

Rappelons que l'ouverture ultime classique est un opérateur générique non paramétrique. Néanmoins, dans les trois approches proposées nous avons sacrifié cette propriété de l'opérateur afin de réduire les effets de masquage et de fuites. Toutefois, les trois stratégies continuent d'être facilement paramétrables par rapport à une application donnée.

Concernant l'approche avec l'information de forme, le paramétrage consiste en un facteur de forme  $(1 + \alpha\psi)$  conçu avec l'information *a priori* des structures à favoriser, en utilisant une fonction de similarité  $\psi$ . La fonction utilise un paramètre additionnel  $\alpha$  qui contrôle l'influence de la similarité entre les formes.

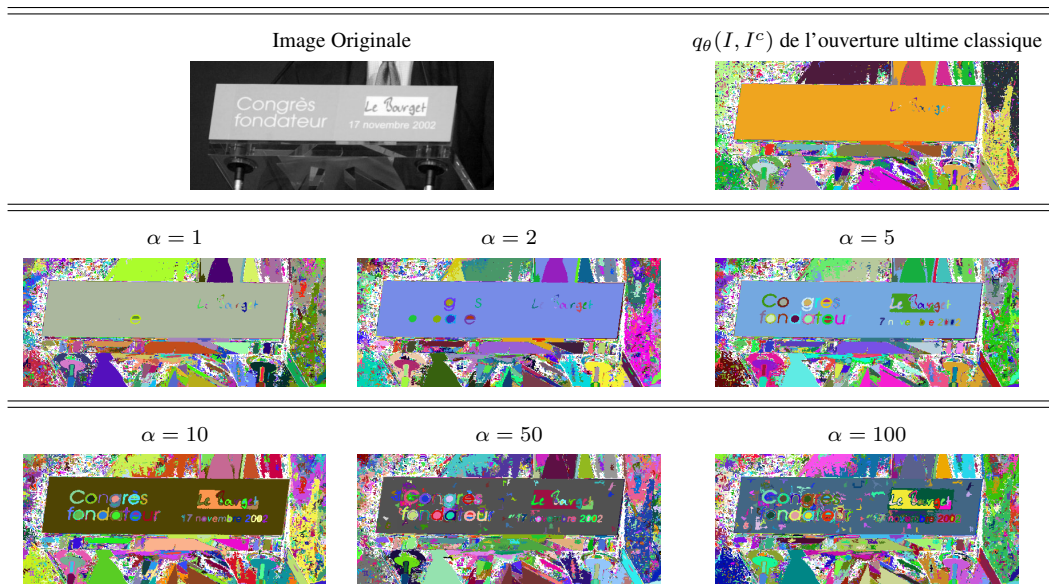


FIGURE 7.21 – Effet du paramètre  $\alpha$ , fonction de forme de l'équation 7.4.  $q_\theta(I, I^c)$  des ouvertures ultimes avec l'information de forme.

La Figure 7.21 illustre la variation de la valeur  $\alpha$  pour la détection de texte enfoui. L'utilisation de l'opérateur classique masque les lettres contenues dans le panneau. En utilisant l'opérateur avec l'information de forme et une valeur  $\alpha = 1$  seulement la lettre *e* est détectée. L'augmentation de la valeur  $\alpha$  favorise la détection des lettres, par exemple avec  $\alpha = 10$  toutes les lettres ont été segmentées. Néanmoins, comme nous pouvons l'observer

sur l'exemple, avec une valeur de  $\alpha \geq 50$  des composantes de bruit sont aussi détectées. Ainsi, avec une valeur élevée de  $\alpha$ , l'information de forme devient prédominante sur l'approche, en réduisant le rôle du résidu pour trouver les structures les plus contrastées. Par ailleurs dans les applications montrées, la valeur de  $\alpha = 9$  a été choisie de manière empirique du fait que c'est la valeur qui produit les meilleurs résultats sur l'ensemble d'images. Cependant, nous proposons d'étudier un paramétrage de  $\alpha$  adapté à chaque image. À propos de l'ouverture ultime hiérarchique, deux paramètres peuvent être choisis ; le niveau de hiérarchie ( $N$ ) et le critère hiérarchique ( $T$ ). Nous avons proposé de ne pas tenir compte du paramètre  $N$  car le critère  $T$  permet d'attendre les niveaux optimaux de l'approche. Nous avons utilisé des critères géométriques comme la surface et des critères couleur basées sur l'écart-type des composantes  $RVB$ . D'autres critères peuvent être utilisés ; cependant, nous avons proposé d'utiliser des critères qui dépendent sur des attributs simples afin de garder le temps de calcul raisonnable comme nous l'avons fait dans l'approche avec la forme.

En ce qui concernait l'approche avec l'accumulation, la valeur de  $\Delta$  est le seul paramètre à définir. Nous proposons d'utiliser des petites valeurs car l'approche favorise la fusion des composantes en augmentant le problème dû aux fuites.

## **Troisième partie**

# **Analyse des Images Couleurs**

# Segmentation des Façades par Bâtiment

## Sommaire

<b>8.1</b>	<b>Introduction</b>	<b>114</b>
<b>8.2</b>	<b>Séparation par bâtiments</b>	<b>116</b>
<b>8.3</b>	<b>Détection du ciel</b>	<b>118</b>
8.3.1	Résultats de la détection du ciel	119
8.3.2	Elimination du ciel	121
<b>8.4</b>	<b>Séparation du rez-de-chaussée et de la chaussée</b>	<b>122</b>
<b>8.5</b>	<b>Discussion</b>	<b>123</b>

## 8.1 Introduction

Nous avons vu dans la section 2.5.3 plusieurs méthodes d'extraction sémantique de façades à partir d'images prises au sol. En analysant ces méthodes en détail, la modélisation de la façade est réalisée en utilisant des images de façade rectifiées, isolées et bien délimitées. Lee and Nevatia en [87] et Wang *et al.* en [170] ont développé des méthodes pour détecter les fenêtres d'une façade d'un bâtiment complètement isolé. Par ailleurs, Müller *et al.* en [114] trouvent des structures architecturales en utilisant l'information mutuelle pour décrire une seule façade, où la façade est manuellement extraite. D'autres approches de découpage manuel de la façade dans la modélisation urbaine sont présentées dans [72, 131].

Cependant, lorsqu'une image d'un bâtiment prise au sol est utilisée pour des méthodes de modélisation d'une seule façade, plusieurs problèmes apparaissent :

1. Des variations de la résolution causées par des effets de la perspective.
2. Plusieurs façades présentes sur la même image.
3. Présence du ciel sur l'image.
4. Présence de la route et des trottoirs.

5. Présence du rez-de-chaussée. Celui-ci est considéré comme un problème car il ne respecte pas la structure répétitive de la façade [114].

Ces problèmes sont illustrés par l'image urbaine de la Figure 8.1.



FIGURE 8.1 –  
Image de façade.

Ce chapitre présente une méthode permettant d'identifier et de segmenter la région image correspondant à une façade de bâtiment au sein d'une image brute (sans hypothèse de filtrage ou de recalage préalable avec d'autres données : des nuages de points ou résultats de reconstructions antérieures type Bati3D). La méthode est complètement automatique à partir de l'image rectifiée. La méthode se résume dans le diagramme de la Figure 8.2. En utilisant une image rectifiée comme entrée, la première étape est la séparation verticale des bâtiments en utilisant un découpage vertical. Ensuite, nous détectons et éliminons les régions du ciel. Finalement, nous divisons par étages et éliminons le rez-de-chaussée car il n'a pas la structure répétitive.

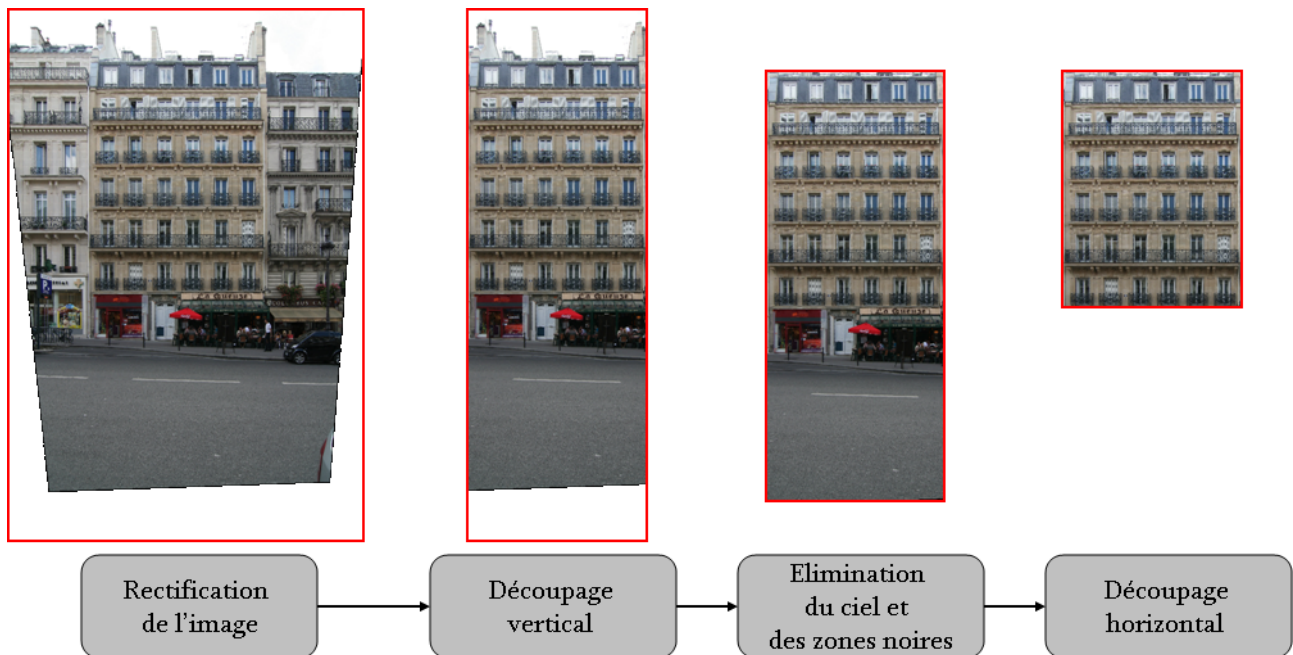


FIGURE 8.2 – Diagramme de la segmentation de façades par bâtiment.

L'étape de rectification est basée sur l'extraction des points de fuite via la transformée de Hough de l'image et une transformation homographique de l'image [61].

## 8.2 Séparation par bâtiments

Nous faisons l'hypothèse que les structures de la façade (fenêtres, portes, balcons) sont alignées horizontalement et verticalement. Pour cette raison, notre approche est basée sur l'analyse de la projection du gradient comme [87] l'a présentée. Nous utilisons les gradients directionnels couleurs pour détecter les différentes structures de la façade. Dans le cas du découpage vertical, nous utilisons seulement le gradient vertical ( $G_v$ ). Ce gradient est généré en utilisant un élément structurant vertical. Avec cette opération, l'approche est plus robuste au bruit généré par les directions perpendiculaires. Ensuite, nous accumulons par colonnes ( $G_v$ ), en créant la projection verticale  $P_v(G_v)$ . Néanmoins, dans les façades texturées typiques de l'architecture haussmannienne, l'accumulation du gradient  $P_v(G_v)$  présente deux problèmes importants : la création de faux pics et de fausses vallées et la réduction de la dynamique entre les deux. Ces problèmes induisent une fausse division de la façade. Pour résoudre ce problème, une opération de filtrage est réalisée avant la projection du gradient. C'est pourquoi nous appliquons une ouverture morphologique directionnelle de taille  $\lambda$ . La direction de l'ouverture est perpendiculaire à la direction du gradient, c'est-à-dire une ouverture horizontale pour le gradient  $G_v$ . La taille idéale de l'ouverture est la taille de la plus petite fenêtre. Cependant, comme cette valeur n'est pas encore connue, une valeur  $\lambda = 10$  a été fixée après plusieurs tests. La sensibilité à ce paramètre n'est pas vraiment importante parce que le filtre éliminera seulement les plus petites structures. La Figure 8.3 présente le gradient vertical  $G_v$  et le gradient filtré  $\widetilde{G}_v$ , ainsi que leurs profils d'accumulation respectifs,  $P_v(G_v)$  et  $P_v(\widetilde{G}_v)$ . On peut observer le débruitage du profil du gradient filtré, ce qui facilite sa segmentation.

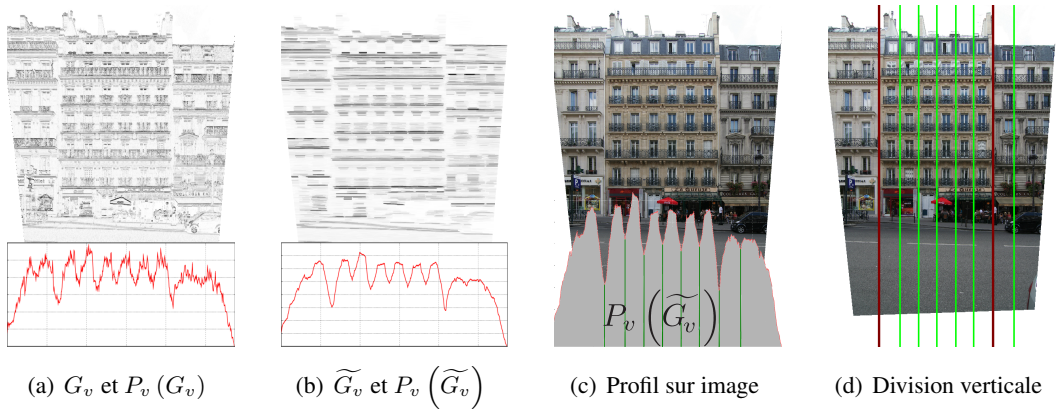


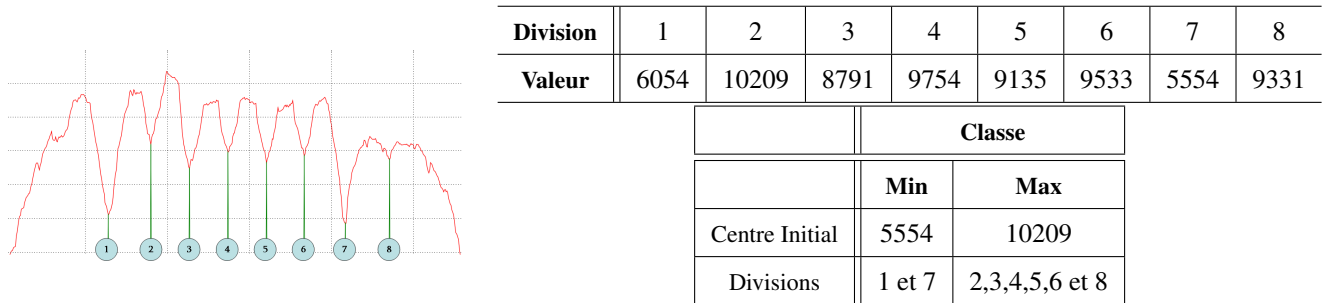
FIGURE 8.3 — (a) Gradient couleur vertical et sa projection. (b) Gradient filtré et sa projection. (c) Projection verticale  $P_v(\widetilde{G}_v)$  sur l'image de départ. (d) Division verticale de la façade.

En outre, en analysant le profil  $P_v(\widetilde{G}_v)$ , on peut voir que les crêtes se trouvent dans la position des fenêtres et les vallées dans celle des murs (voir Figure 8.3(c)). Ensuite, le profil est filtré en utilisant un Filtre Alterné Séquentiel (FAS) afin de faciliter la détection des maxima et minima les plus représentatifs. Nous utilisons comme taille de



filtre la même taille de l'ouverture (10 pixels). Finalement, pour trouver la division des façades, nous utilisons la ligne de partage des eaux par marqueurs [15] du profil original inversé et comme marqueurs les maxima du profil filtré  $FAS \left( P_v \left( \widetilde{G}_v \right) \right)$ . Le résultat de la division verticale est illustré sur la Figure 8.3(d). La sélection de la taille du  $FAS$  affecte le nombre de divisions verticales détectées. La diminution de la valeur de 10 produira une sur-détection et vice-versa. Sur la base d'images, la valeur de 10 génère dans la plupart des cas une détection de toutes les divisions, néanmoins dans certains cas, une ou deux divisions sont manquées. La non-détection de quelques divisions ne pose pas de problème à la segmentation verticale car les divisions qui séparent plusieurs bâtiments seront toujours détectées.

Parmi les divisions trouvées, nous avons deux types de division : intra-façades et inter-façades. Comme nous pouvons le voir dans la Figure 8.3(c), les divisions inter-façades sont les vallées les plus profondes dans le profil et elles ont donc une valeur plus basse sur le profil. Ce comportement est dû au fait qu'il n'existe pas de structures horizontales sur la ligne de séparation entre deux bâtiments, tandis que dans la façade, on trouve des balcons filants, des ornements et des dalles entre étages, qui produisent plus de réponses sur le gradient et donc, une valeur plus élevée d'accumulation. Par conséquent, afin de classer les deux types de division, l'algorithme de  $k$ -means,  $k = 2$ , est utilisé. Nous prenons comme centres initiaux des deux groupes le minimum et le maximum parmi les valeurs des divisions. Enfin, les divisions qui appartiennent au groupe du minimum sont les séparations entre les façades (Figure 8.3(d) : lignes rouges). La Figure 8.4 montre le calcul du  $k$ -means sur le profil de cet exemple où deux divisions, 1 et 7, sont trouvées comme les lignes qui séparent les façades.

FIGURE 8.4 – Sélection des divisions par  $k$ -means

L'algorithme de  $k$ -means sur les valeurs du profil dans la division trouve de manière précise et automatique les séparations entre façades. Cependant, une dernière considération doit être prise en compte. Dans l'architecture haussmannienne, les façades très larges ont certaines divisions intra-façade avec des valeurs du profil trop basses, qui seront donc classées comme inter-façade. Ces divisions se trouvent principalement sur l'endroit du mur où est situé le système de la gouttière. Voyons la Figure 8.5 où, en utilisant la méthode décrite, une fausse séparation a été détectée. C'est pourquoi une dernière contrainte de distance entre séparations a été ajoutée. Si la distance entre deux séparations est inférieure à un seuil donné, la séparation avec l'évaluation la plus haute sera éliminée. Dans

notre cas, nous avons choisi le seuil du tiers de la largeur de l'image.

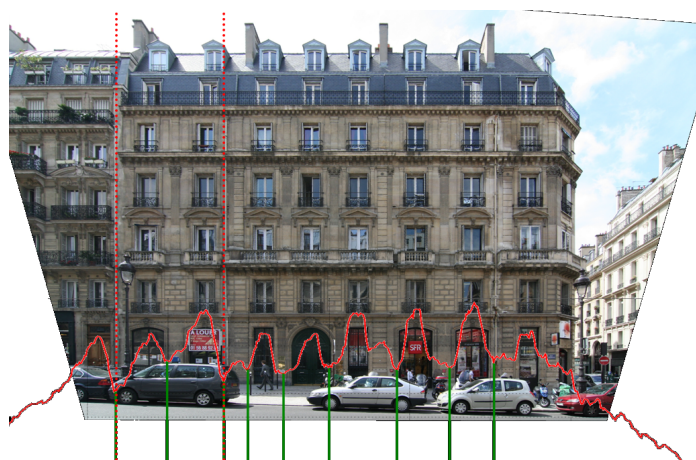


FIGURE 8.5 – Fausse détection inter-façade.

### 8.3 Détection du ciel

La détection du ciel est un problème courant de l'extraction de sémantique dans les images, pour lequel le nombre de brevets a été très important. Cette détection a été utilisée pour classer les scènes d'intérieur-extérieur [96], pour connaître l'orientation des images [51] ou comme marqueur pour détecter la sémantique des images urbaines [170]. Nous proposons une méthode basée sur l'information couleur et un filtre connexe. La méthode est conçue pour une détection de ciel d'images urbaines. Ces images ont des caractéristiques du ciel particulières car elles sont prises pendant la journée et avec une orientation connue. Nous nous sommes basés sur deux approches comme point de départ de notre méthode : la combinaison d'une segmentation en composantes connexes et un marqueur de couleur [134], et la détection d'une région graine pour caractériser la couleur du ciel et la sélection d'autres régions candidates [51].

Comme résultat de la correction de perspective, certaines zones noires apparaissent dans l'image. Ces zones, en particulier les zones en haut et en bas de l'image, produisent plusieurs problèmes dans la détection du ciel et la séparation horizontale (élimination du rez-de-chaussée, section 8.4). Ainsi, une méthode simple pour enlever ces zones noires est mise en œuvre. Elle consiste à éliminer tous les segments horizontaux noirs de plus de un quart de la largeur de l'image après la séparation verticale. En ce qui concerne les zones noires verticales, elles sont éliminées lors de la séparation verticale des bâtiments (voir Figure 8.2, page 115). Notre méthode est décrite dans le diagramme de la Figure 8.6 par les étapes suivantes :

- **Segmentation** : tout d'abord, l'image est segmentée en utilisant l'opérateur d'ouverture ultime [14] de l'image gradient Lab inversée. Le choix de l'espace de couleur Lab permet une bonne segmentation du ciel grâce à la

propriété perceptuellement uniforme de l'espace (proximité de la vision humaine).

- **Marqueur bleu** : nous calculons une image marqueur pour déterminer les pixels qui sont potentiellement du ciel. Nous faisons l'hypothèse que le ciel est normalement de couleur bleue pour créer le marqueur. Donc, un pixel est « *bleu* », si la composante bleue est plus élevée que les composantes rouge et verte  $B > \max(R, V)$ . Si le pixel est « *bleu* », nous assignons 255 au pixel du marqueur, autrement la valeur de la composante  $B$  est attribuée au pixel. En outre, les images étant prises en plein jour, le ciel présente une luminance élevée. Ainsi, une valeur de 140 est choisie comme la luminance minimale acceptée dans le marqueur. Cette valeur est le seuil utilisé pour toutes les expériences dans l'ensemble des tests de la méthode.
- **Régions candidates** : en utilisant la segmentation et le marqueur bleu, nous voulons extraire les composantes connexes, qui sont potentiellement du ciel. Une image mosaïque est produite, pour laquelle la moyenne du marqueur est la couleur associée à chaque région. La mosaïque indique le pourcentage de bleu que contient une région. La première sélection des régions candidates se fait par l'élimination des régions de faible valeur dans l'image mosaïque, pour tous les tests nous utilisons un seuil de 90% du marqueur bleu, c'est-à-dire la valeur égale à 230. Ensuite, un critère de localisation est employé en sélectionnant les régions qui ont au moins un point sur le bord de l'image.
- **Région graine** : Cette région est celle qui statistiquement caractérise la couleur du ciel sur l'image. Pour la choisir, trois conditions sont demandées : elle est en haut de l'image, elle est la plus grande région parmi les régions candidates et elle a la plus grande valeur de luminance. Cependant, parfois la plus grande région n'a pas la plus haute valeur de luminance. Alors, les deux valeurs sont multipliées (luminance et surface) afin de sélectionner la région graine. Finalement, la région graine est modélisée en utilisant une distribution gaussienne pour les trois composantes couleurs  $R, V, B$ .
- **Sélection du ciel** : Parmi les régions candidates, nous sélectionnons toutes les régions compatibles avec le modèle de couleur calculé avec le critère de  $|\mu_{\text{modèle}} - \mu_{\text{région}}| < \sigma_{\text{modèle}}$ . Ainsi, 95% de la distribution est sélectionnée.

La méthode est relativement simple pour détecter le ciel. Des approches plus structurées [51, 96] sont disponibles pour améliorer la robustesse. En outre, ces méthodes cherchent parmi toutes les régions de la segmentation pour sélectionner les régions candidates. Dans notre cas, ce n'est pas la meilleure solution, car la réflexion sur les fenêtres produit beaucoup de faux positifs en milieu urbain.

### 8.3.1 Résultats de la détection du ciel

Pour tester la méthode de détection du ciel, une base de données d'images urbaines a été utilisée (490 images avec du ciel et 52 sans ciel). Globalement, le taux de détection est de 81,56%. La Figure 8.7 montre quelques exemples de résultats expérimentaux démontrant la performance de la méthode. Nous illustrons les cas caractéristiques du ciel dans les zones urbaines où le ciel est correctement détecté. La Figure 8.7(a) est un cas d'une gradation blanc-

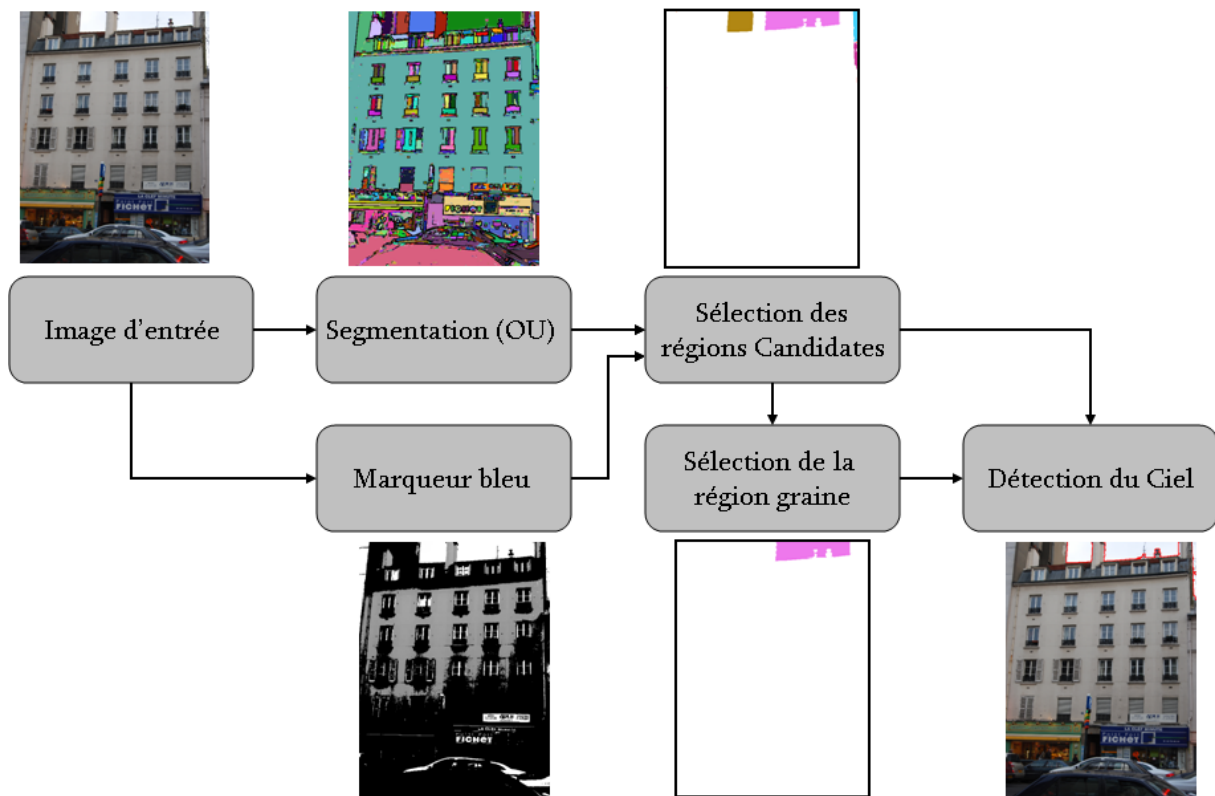


FIGURE 8.6 – Diagramme de la détection du ciel.

bleu des couleurs de gauche à droite, la Figure 8.7(b) montre un ciel divisé par un obstacle, la Figure 8.7(c) illustre une petite région du ciel, et dans la Figure 8.7(d) un ciel nuageux est présenté.

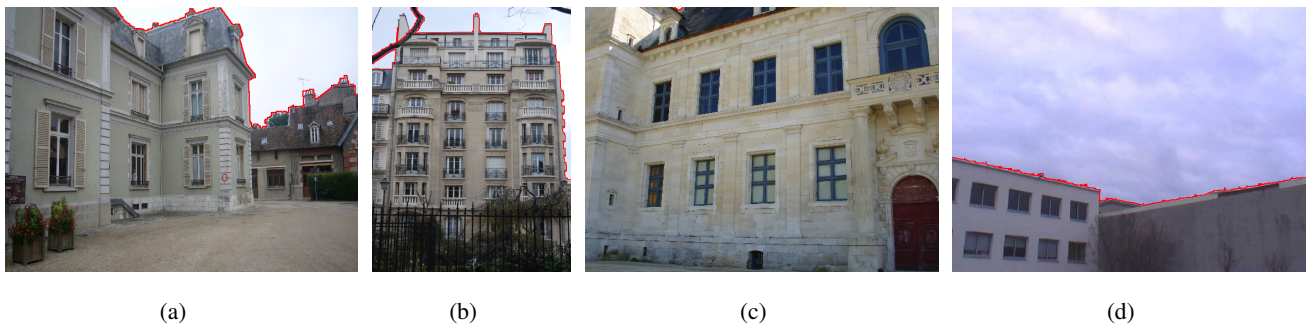


FIGURE 8.7 – Résultats de détection du ciel.

La Figure 8.8 montre les cas typiques d'échec. Une analyse détaillée des résultats utilisant la matrice de confusion est illustrée sur la Table 8.1. Nous avons séparé les résultats en deux catégories, avec et sans ciel, et nous avons considéré deux groupes, bonne et mauvaise détection. Pour les images avec ciel, la mauvaise détection a été

subdivisée en :

- *Sur-détection* : Détection de toutes les régions du ciel et d'autres non ciel.
- *Sous-détection* : Détection de quelques régions du ciel.
- *Sur / Sous -détection* : Détection de quelques régions ciel et d'autres non ciel.

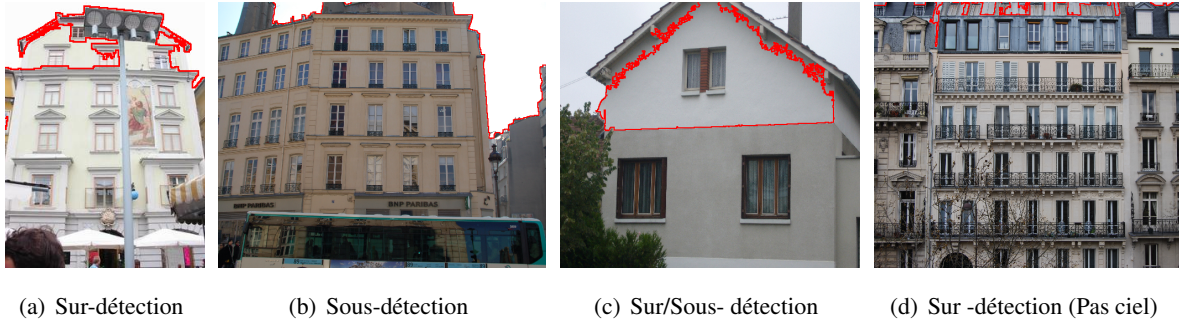


FIGURE 8.8 – Cas d'échec de la détection du ciel.

Pour les images sans ciel, seule la sur-détection est considérée comme mauvaise détection (voir Figure 8.8(d)). La sur-détection est principalement produite par les reflets de fenêtres et par des structures similaires à la couleur du ciel (les toits bleuâtres). Ces structures produisent une mauvaise sélection de la région graine ou des régions candidates. Elle est aussi générée par un faible contraste entre le ciel et le bâtiment, comme l'illustre la Figure 8.8(a), en affectant l'étape de segmentation. La sous-détection est causée par deux facteurs : la mauvaise sélection de la région graine et/ ou le fait que la région graine ne soit pas représentative de toutes les régions du ciel dans l'image (voir Figure 8.8(b)). Nous pouvons constater que le principal problème est dû à une mauvaise détection sur les images sans ciel, en réduisant la performance de notre méthode.

	Bonne Détection	Mauvaise Détection			Total
		Sur-Detec.	Sous-Detec.	Sur/Sous Detec.	
<b>Ciel</b>	421 (85,91%)	21 (4,28%)	28 (5,71%)	20 (4,08%)	490 (100%)
<b>Non Ciel</b>	21 (40,38%)	31 (59,62%)			52 (100%)
<b>Total</b>	442 (81,56%)	100 (19,44%)			542 (100%)

TABLE 8.1 – Matrice de Confusion de la détection du ciel.

### 8.3.2 Elimination du ciel

Dans l'objectif d'une extraction de la façade, le ciel est éliminé après sa détection. L'élimination se fait d'abord horizontalement et ensuite verticalement si cela est nécessaire. La ligne de séparation horizontale est l'endroit où il y a le plus grand nombre de pixels au bord des régions du ciel. De même, une procédure similaire est appliquée



sur le sens vertical. La Figure 8.9 illustre les lignes d'élimination du ciel. Notez que les cheminées sont éliminées car la ligne principale est sur le toit.

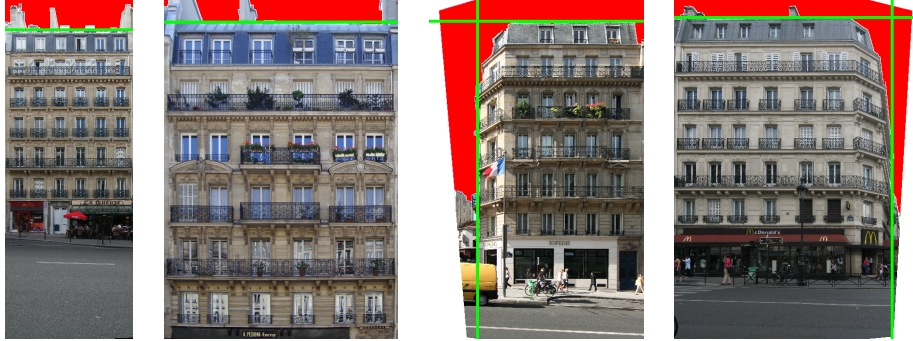


FIGURE 8.9 – Détection et lignes d'élimination du ciel.

## 8.4 Séparation du rez-de-chaussée et de la chaussée

Comme nous pouvons l'observer dans l'étape 3 de la Figure 8.2, le bâtiment a été bien découpé verticalement et au niveau du ciel. Par contre, il reste l'information de la chaussée et du rez-de-chaussée du bâtiment. Il est donc nécessaire de découper horizontalement l'image et d'éliminer les sections qui ne sont pas des étages. En utilisant la même procédure de la projection de gradient, présentée dans la section 8.2, une première division horizontale est mise en œuvre. Dans ce cas le gradient couleur horizontal ( $G_h$ ) est produit par un élément structurant horizontal (voir Figure 8.10(a)). Ensuite, le gradient est filtré avec une ouverture verticale de taille  $\lambda = 10$ , en générant  $\widetilde{G}_h$  (voir Figure 8.10(b)). Nous projetons  $\widetilde{G}_h$  par lignes afin de créer le profil  $P_h(\widetilde{G}_h)$  (voir Figure 8.10(c)). Le profil est filtré par un  $FAS$  comme dans le cas vertical, et un opérateur  $h$ -Minima [150]. L'opérateur  $h$ -Minima élimine les minima qui ont une profondeur inférieure ou égale à un seuil  $h$ . Il utilise une reconstruction morphologique par érosion ( $R_f^\varepsilon$ ) où le marqueur est une image à laquelle on a additionné une constante  $h$  (voir Equation 8.1).

$$\text{HMin}(f) = R_f^\varepsilon(f + h) \quad (8.1)$$

Dans notre cas, la valeur  $h$  a été choisie comme le dixième de la distance entre le maximum et le minimum du profil. Cet opérateur a été ajouté afin d'éviter les divisions sur les régions homogènes, telles que les routes. Finalement, nous utilisons la ligne de partage des eaux du profil inversé filtré pour trouver les divisions horizontales.

Pour l'instant nous avons trouvé les divisions horizontales de la façade grâce au gradient produit par les fenêtres (Figure 8.10(b)). C'est pourquoi la séparation entre étages n'est pas positionnée de manière précise. Pour résoudre ce problème, nous déplaçons par étage la division sur le maximum de la projection du gradient vertical ( $\widetilde{G}_v$ ), ce qui donne comme résultat la Figure 8.10(d).

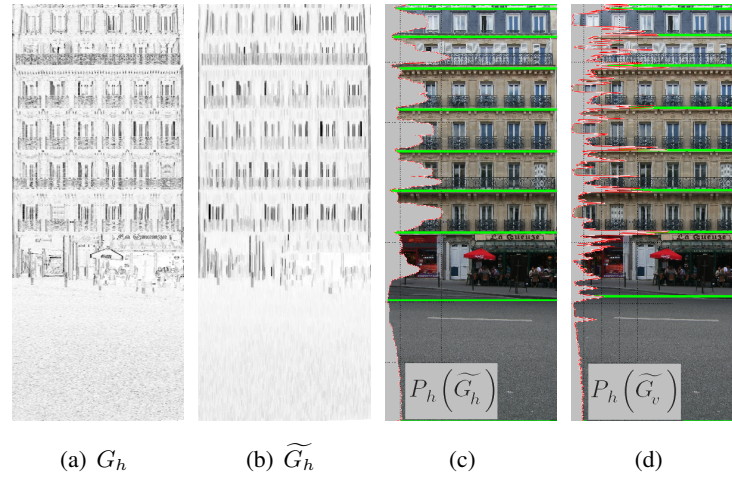


FIGURE 8.10 – (a) Gradient couleur horizontal et (b) gradient filtré. (c) Projection horizontale de  $\widetilde{G}_h$  et division horizontale de façade, (d) Projection horizontale de  $\widetilde{G}_v$  et division horizontale raffinée de façade.

Dans le but d'éliminer les régions qui ne sont pas des étages, le rez-de-chaussée et la route, nous analysons étage par étage du bas vers le haut. Dans l'exemple que nous sommes en train d'illustrer deux sections devront être supprimées. Afin de déterminer si une section est un étage de la façade, nous comparons la projection normalisée locale du gradient vertical par étages  $P_v \left( \widetilde{G}_{v_{\text{étage}}} \right)$  avec projection normalisée totale du gradient vertical de toute l'image  $P_v \left( \widetilde{G}_v \right)$ . La Figure 8.11 illustre cette analyse. Notez que la différence des projections est plus élevée quand la section analysée ne correspond pas à un étage de la façade (section 7 rez-de-chaussée). Pour mesurer cette différence, nous calculons l'écart-type de la différence entre projections. Avec l'algorithme *k-means* de la différence, deux classes sont établies : étages et non-étages. Nous éliminons les sections du bas vers le haut jusqu'à ce que nous trouvions une section de la classe étage. Le résultat final de la segmentation de la façade par bâtiment peut être observé dans la Figure 8.2, page 115.

## 8.5 Discussion

Dans ce chapitre nous avons décrit une méthode pour la segmentation de la façade par bâtiment. A partir d'une image de façade prise au sol rectifiée, la méthode extrait automatiquement une image de façade délimitée. La méthode a trois étapes principales : le découpage vertical, l'élimination de zones noires et du ciel, et l'élimination de zones non étages en incluant le rez-de-chaussée. Chacune de ces étapes utilise plusieurs filtres morphologiques afin d'augmenter leur robustesse.

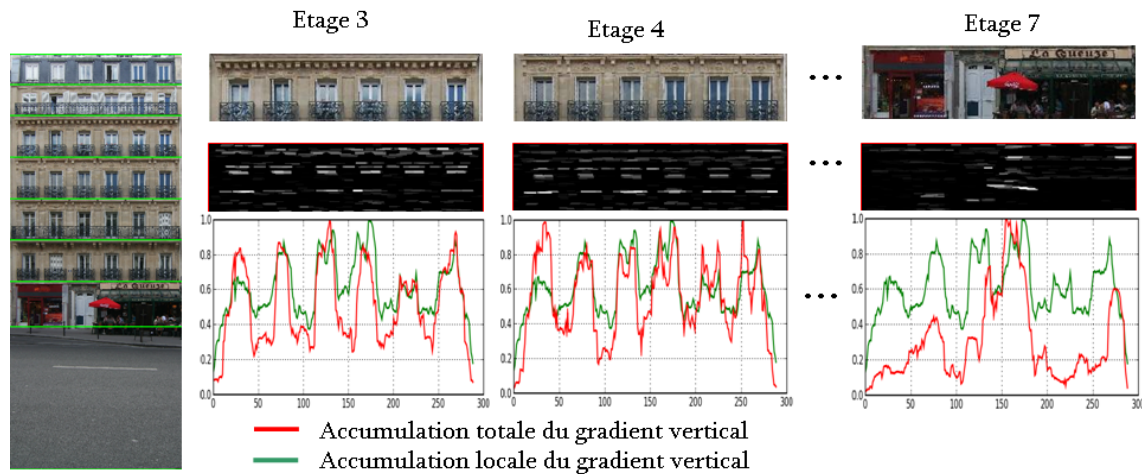


FIGURE 8.11 – Analyse horizontale par étages.



FIGURE 8.12 – Ligne inter-façade occlue pour un arbre.

Concernant le découpage vertical, il permet de séparer les différentes façades d'une image. La méthode s'appuie sur la segmentation de la projection verticale du gradient vertical. Le découpage vertical utilise également l'algorithme de *k-means* pour classer les divisions inter et intra façade. Le fait d'utiliser l'algorithme de *k-means* implique qu'il y ait au moins une division verticale sur l'image que nous sommes en train d'analyser. Cependant si les centres des deux classes trouvées sont trop proches, nous pourrions nous rendre compte que la séparation inter-façade n'est pas nécessaire. Par ailleurs, comme l'algorithme de *k-means* est appliqué sur la valeur de la projection sur la division, les divisions inter-façades doivent être les plus dégagées possible. Par exemple, la Figure 8.12 montre une occlusion par un arbre dans la division de la ligne inter-façade qui produit une mauvaise séparation des bâtiments.

La détection du ciel sélectionne automatiquement les régions du ciel dans l'image. Le principal problème de la méthode est la détection du ciel quand l'image n'en a pas. Néanmoins, comme la méthode globale est complètement modulaire, un utilisateur pourrait enlever cette procédure dans le cas où il n'en aurait pas besoin. Par ailleurs, une sous-détection du ciel n'est pas considéré un problème trop important, lorsque la méthode détecte la plus grande région horizontale du ciel. En effet, en détectant les régions principales du ciel, la ligne d'élimination passera au bon endroit.

L'élimination des sections qui ne sont pas des étages se fait en utilisant une division horizontale de l'image. Cette division est réalisée de manière analogue à la division verticale en utilisant la projection horizontale du gradient



horizontal. Ainsi, une première division horizontale est faite, et ensuite raffinée sur les pics de la projection horizontale du gradient vertical. Chaque division est analysée en utilisant la projection verticale du gradient vertical et en comparant avec la projection globale. Les sections sont éliminées de manière itérative du bas vers le haut. De la même façon, nous utilisons l'algorithme de *k-means* pour classer étage et non-étage, en supposant que les deux classes se trouvent sur l'image. Si cela n'est pas le cas, la méthode fournira de toute façon un résultat, qui pourrait être validé par la suite, en fonction de la différence entre les deux classes. Nous avons testé l'ensemble de la méthode sur la base de données du projet (30 images de façade prises au sol, voir section 1.3), en ayant un taux de succès du 100%. La Figure 8.13 montre quelques résultats de la méthode. Les autres résultats sont disponibles sur le site web suivant : [cmm.ensmp.fr/~hernandez/facade\\_en.html](http://cmm.ensmp.fr/~hernandez/facade_en.html). Cependant, des tests sur une base de données plus large seront nécessaires pour la validation de cette méthode.



(a) Images d'entrée



(b) Images de sortie

FIGURE 8.13 – Résultats de la segmentation de façades par bâtiments.

# Extraction de la Sémantique d'une Façade

## Sommaire

<b>9.1</b>	<b>Introduction</b>	<b>126</b>
<b>9.2</b>	<b>Division de la façade</b>	<b>126</b>
9.2.1	Segmentation en étages et détection de la dalle	127
9.2.2	Division en travées	128
<b>9.3</b>	<b>Détection de fenêtres et de balcons</b>	<b>129</b>
9.3.1	Localisation de fenêtres	131
9.3.2	Détection de balcons	132
<b>9.4</b>	<b>Modélisation procédurale</b>	<b>134</b>
<b>9.5</b>	<b>Discussion</b>	<b>135</b>

## 9.1 Introduction

Dans ce chapitre nous présentons l'analyse d'une image de façade délimitée, issue de la méthode décrite par le chapitre précédent. Cette analyse consiste en l'extraction de l'information sémantique telle que le nombre d'étages, le nombre de colonnes, les fenêtres, pour fournir de manière automatique des informations importantes dans une modélisation procédurale, basée sur des règles.

## 9.2 Division de la façade

Une description complète d'une façade haussmannienne a été présentée par Janil *et al.* en [77]. Cette description sera largement utilisée, dans ce chapitre, afin d'établir les différents éléments sémantiques à extraire dans l'image de la façade. La description est faite de manière hiérarchique depuis une façade, jusqu'à ses ornements. La Figure 9.1 présente la subdivision d'une façade haussmannienne. Ce diagramme est une version simplifiée de la description de la façade car il n'illustre que la décomposition jusqu'aux travées. La façade est d'abord divisée en étages

avec la dalle incluse ; ensuite, la dalle est séparée de l'étage. La dalle est classée comme dalle de balcon filant ou dalle de balcon individuel. En fin, chaque étage est divisé en travées.

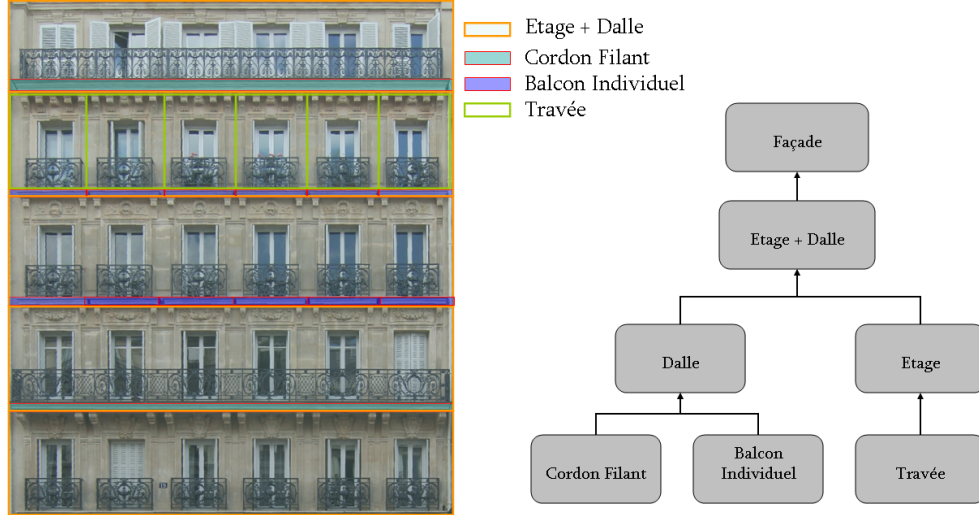


FIGURE 9.1 – Diagramme simplifié de subdivision d'une façade haussmannienne.

A la différence de la division présentée par [77], nous proposons de diviser la façade en étages et de détecter la dalle mais sans distinguer les deux types de dalle. Ensuite, la division en travées est réalisée en utilisant une division en colonnes de la façade complète pour rendre plus robuste la méthode.

### 9.2.1 Segmentation en étages et détection de la dalle

La première étape de l'extraction sémantique correspond à la division en étages de la façade. Cette étape a déjà été décrite dans le processus de séparation du rez-de-chaussée et de la chaussée (voir section 8.4). La division horizontale trouvée est localisée sur le maximum par étage de la projection horizontale de  $\widetilde{G}_v$ . Cependant, la division est normalement localisée dans la partie haute de la dalle. En regardant la subdivision proposée par [77], cette division devrait se trouver dans la partie basse.

Comme nous pouvons l'observer, la localisation de la dalle est proche de la division horizontale de la façade déjà détectée. Ainsi, à partir de la division horizontale, nous déterminons une zone d'analyse d'un dixième de la hauteur moyenne des étages<sup>1</sup>. A l'intérieur de cette zone, nous extrairons les deux maxima de la projection horizontale du gradient  $\widetilde{G}_v$ . Les deux maxima seront conservés si leurs valeurs sont supérieures à la moyenne entre le maximum et le minimum du profil dans la zone d'analyse. Le résultat de cette procédure est illustré sur la Figure 9.2 où les dalles sont automatiquement détectées.

1. La hauteur moyenne des étages est calculée en utilisant la division horizontale

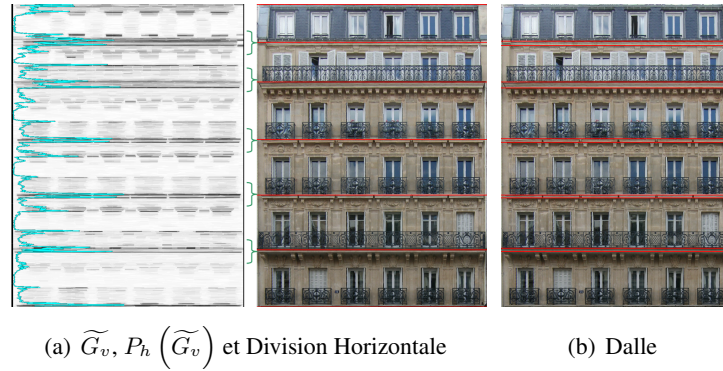


FIGURE 9.2 – (a) Projection horizontale de  $\widetilde{G}_v$  et division horizontale raffinée de façade. (b) Détection de la dalle.

### 9.2.2 Division en travées

La division en travées est réalisée sur la division en colonnes de la façade. Cette division est un cas particulier de la division verticale, décrite par la section 8.2, pour séparer les bâtiments. La méthode est basée sur l'accumulation verticale du gradient vertical ( $G_v$ ). Dans le cas de la séparation des bâtiments, une ouverture morphologique directionnelle est utilisée afin d'éliminer le bruit de l'image gradient. Le profil généré présente une différence très importante dans la dynamique, grâce aux balcons filants, aux ornements et aux dalles entre étages, qui permet d'identifier les vallées les plus importantes où se trouvent les séparations. Par contre, à l'intérieur de la façade, ces éléments peuvent empêcher la bonne détection de toutes les divisions verticales. C'est pourquoi, l'image originale est filtrée avec une fermeture verticale, de taille d'un dixième de la hauteur moyenne des étages. Le gradient est aussi filtré, en éliminant les petites structures (ouverture de taille  $\lambda = 10$ ) et les grandes structures (chapeau haut de forme de l'ouverture, de taille un quart de la largeur de l'image). La Figure 9.3 illustre deux exemples de la projection et du filtrage du gradient nécessaire pour la localisation de la division verticale. La Figure illustre également le cas de la projection du gradient de l'image originale et de la projection du gradient de l'image filtrée. Pour le premier exemple, le fait de filtrer l'image d'entrée ne change pas le résultat vis-à-vis du nombre de maxima et minima dans le profil final (voir Figures 9.3(a) et 9.3(b)). Par contre, dans le deuxième exemple l'amélioration est concluante. En effet, en ajoutant la fermeture de l'image d'entrée, le profil a les maxima et les minima nécessaires pour la bonne division verticale de la façade (voir Figures 9.3(c) et 9.3(d)). La fermeture réduit donc les détails des façades et les balcons caractéristiques de l'architecture haussmannienne, en rendant la méthode plus robuste.

Ainsi, la méthode de division en travées est résumée par les étapes suivantes :

1. Filtrage de l'image couleur par une fermeture de taille d'un dixième de la hauteur moyenne des étages.
2. Extraction du gradient horizontal avec un élément structurant vertical.
3. Filtrage de l'image gradient par une ouverture de taille 10 et par un chapeau haut de forme d'une ouverture

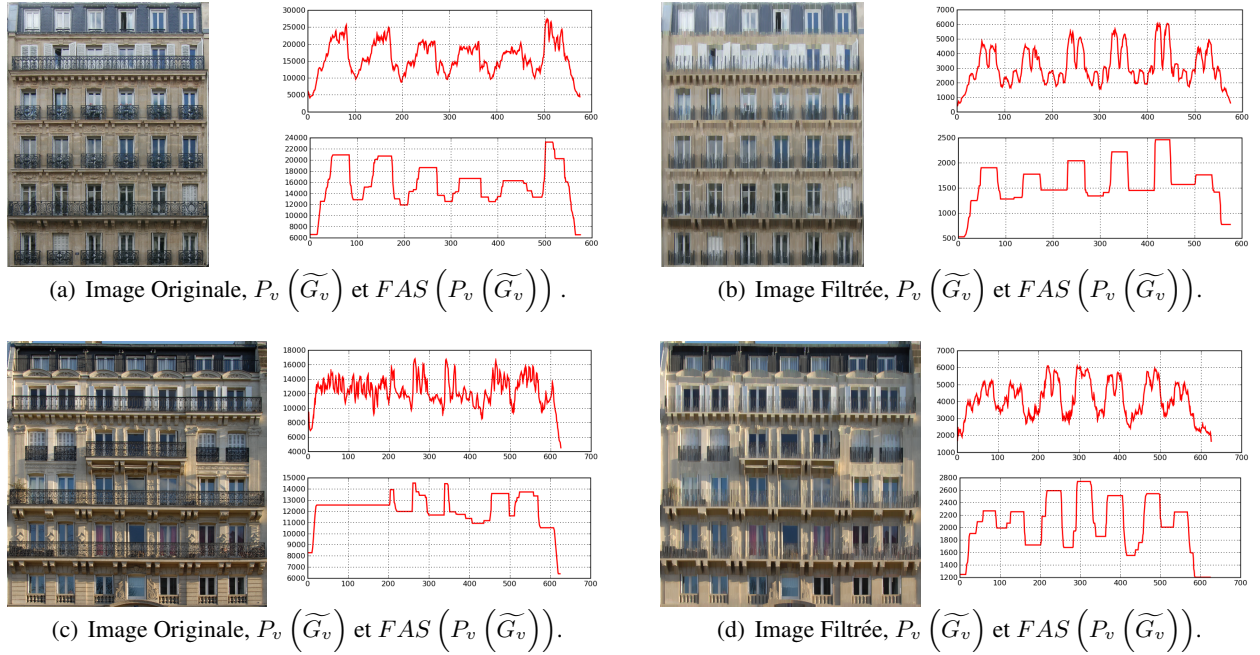


FIGURE 9.3 – Effets du filtrage vertical sur le profil.

d'un quart de la largeur de l'image.

4. Génération du profil avec l'accumulation verticale du gradient.
5. Filtrage de profil avec un  $FAS$ .
6. Ligne de partage des eaux sur profil inversé, contraint aux maxima du profil filtré.

Jusqu'à présent nous avons fixé à 10 pixels la taille de  $FAS$  ; cependant avec cette valeur la division verticale peut être incomplète. En profitant du fait que le traitement est réalisé en une dimension sur la projection du gradient, nous proposons de tester une série de tailles de filtre. Ainsi, nous trouverons la taille optimale du filtre sur chaque façade, comme celle qui minimise l'écart-type de la largeur des travées après chaque réalisation.

La Figure 9.4 illustre quelques exemples de la division horizontale et verticale de la façade dont nous pouvons observer la bonne détection et la séparation des dalles et des travées.

### 9.3 Détection de fenêtres et de balcons

Dans la description présentée par Janil *et al.* en [77], les travées ont été aussi subdivisées en tous les éléments qui les composent (voir la Figure 9.5). La décomposition est réalisée d'abord en linteau et mur, et elle continue de





FIGURE 9.4 – Division de la façade en étages + dalle et en travées.

manière hiérarchique jusqu'aux consoles, fenêtres, trumeaux, etc.

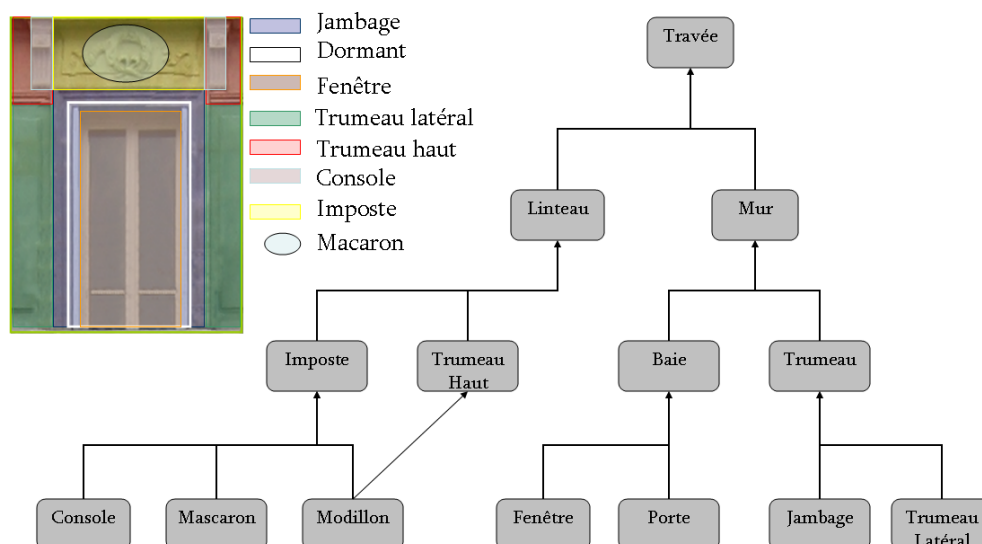


FIGURE 9.5 – Diagramme simplifié de subdivision d'une travée.

Comme nous pouvons l'observer dans la Figure 9.5, la description d'une travée est réalisée sans le décor (en pierre ou en fer forgé) des balcons. Cependant, cette information est présente dans les travées issues de notre division de façade. Ainsi, nous centrons notre analyse de la travée sur la détection de la fenêtre et du décor du balcon. Pour nous, la localisation fenêtre inclut le dormant, car la séparation entre ces deux structures est trop fine dans l'image. D'ailleurs, avec la localisation du décor du balcon, nous pouvons déterminer l'information de cordon filant ou balcon individuel afin de déterminer le type de dalle. De plus, avec la largeur du balcon nous pouvons calculer l'épaisseur du jambage et des trumeaux.

### 9.3.1 Localisation de fenêtres

La détection de fenêtres se fait à l'aide de la division préalable car nous supposons qu'il n'y a qu'une fenêtre par travée. Analysons d'abord la localisation verticale. Nous supposons que la bordure verticale de la fenêtre présente une forte réponse dans le gradient horizontal ( $G_h$ ). Si toutes les fenêtres sont alignées, le gradient générera des pics importants après un processus d'accumulation  $P_v(\widetilde{G}_h)$ . D'ailleurs, nous projetons le gradient filtré  $\widetilde{G}_h$  afin d'éviter les problèmes de bruit expliqués précédemment. Ensuite, en analysant les maxima et minima du profil de la division en travées ( $FAS(P_v(\widetilde{G}_v))$ ), nous avons constaté que les minima passent à travers le mur et les maxima à travers les fenêtres. En utilisant cette information, nous appliquons une ligne de partage des eaux avec contraintes sur le profil  $P_v(\widetilde{G}_h)$ . La Figure 9.6 présente chacune des étapes dans la détection de la limite verticale des fenêtres. La méthode permet d'estimer de manière globale la localisation de fenêtres ; cependant les effets de la perspective deviennent plus visibles. Comme nous pouvons l'observer, les lignes sont bien placées dans les étages inférieurs et centraux. Par contre, dans les étages supérieurs et latéraux, nous observons le décalage de la bordure fenêtre et ligne. Une stratégie pour résoudre ce problème est de raffiner de manière locale et par travée les limites verticales de chaque fenêtre.

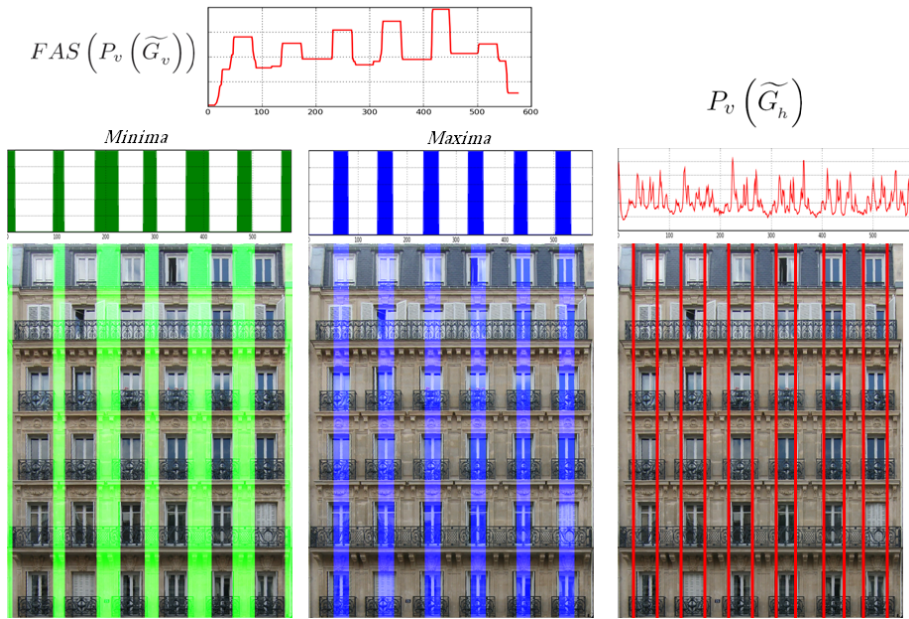


FIGURE 9.6 – Localisation de la bordure verticale des fenêtres.

Par ailleurs la localisation horizontale est basée sur les mêmes principes que la division verticale ; néanmoins, nous ne détectons que les lignes supérieures des fenêtres, car nous supposons que la limite inférieure de la fenêtre est la dalle. Ainsi, nous utilisons pour chaque étage la ligne de partage des eaux avec contraintes, sur la projection  $P_h(\widetilde{G}_v)$  avec les marqueurs des minima et des maxima de  $FAS(P_h(\widetilde{G}_h))$ . Le problème de la perspective est

moins fort à l'horizontal grâce à la position d'acquisition des photos. Finalement, nous combinons les localisations verticales et horizontales afin d'estimer les positions des fenêtres. Des exemples de la méthode sont illustrés sur la Figure 9.7.

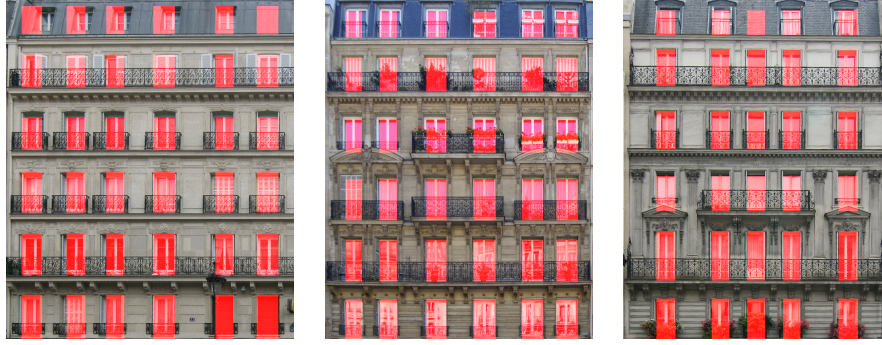


FIGURE 9.7 – Localisation de fenêtres.

### 9.3.2 Détection de balcons

Dans notre étude nous ne nous sommes intéressés qu'à la détection des balcons en fer forgé. La détection est réalisée sur l'image de luminance de la façade et en utilisant la position horizontale de la fenêtre, car nous supposons que les balcons sont des structures sombres toujours en face des fenêtres. L'image est d'abord filtrée à l'aide d'une ouverture verticale et horizontale pour relier les possibles déconnexions de la rambarde et des barreaux en fer forgé. Ensuite nous remplissons l'intérieur des balcons, en utilisant une reconstruction morphologique par dilatation  $R_f^\delta(g)$ , afin d'avoir des structures sombres plus homogènes. L'image filtrée par les ouvertures est utilisée comme référence  $f$ . Le marqueur  $g$  est la même image de référence, en mettant à zéro la partie basse des étages, de la dalle jusqu'à mi-hauteur des fenêtres. Un exemple des images intermédiaires de la procédure est illustré sur la Figure 9.8. Nous pouvons observer que dans l'image reconstruite les balcons sont les structures les plus sombres.

La méthode de détection de balcons utilise les méthodes de projection présentées dans les sections précédentes. Nous avons d'abord localisé les balcons à l'horizontal et ensuite, nous détectons sur la région horizontale trouvée, les séparations verticales. La projection par lignes est réalisée sur l'inverse de l'image  $R_f^\delta(g)$ . Ainsi, les balcons seront des "bosses". Cette projection est réalisée étage par étage séparément, et nous ne projetons que la moitié inférieure de la localisation horizontale des fenêtres. La projection nous permet de détecter la limite supérieure des balcons en étages car nous supposons la dalle comme la limite inférieure. La limite supérieure est détectée à l'aide de la ligne de partage des eaux avec contraintes sur gradient du profil et comme contraintes le maximum et le minimum du profil. Une fois que la détection horizontale est réalisée, nous projetons par colonnes la région des balcons. Nous appliquons un seuil au profil généré. La valeur du seuil est le point au milieu entre la moyenne des



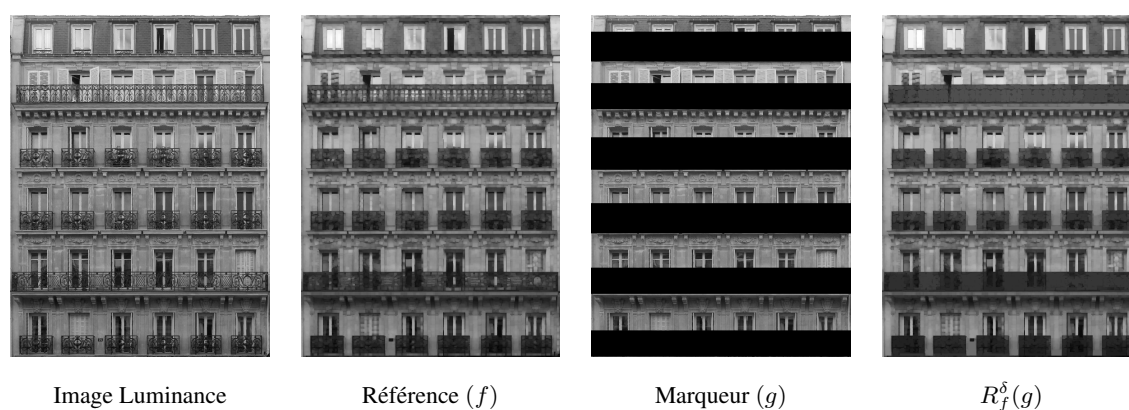


FIGURE 9.8 – Adéquation de l'image de luminance pour renforcer la couleur sombre des balcons.

maxima et la moyenne des minima de profil projeté verticalement. Cette procédure est résumée sur le Diagramme de la Figure 9.9(a). Les résultats de la détection horizontale et verticale de cet exemple sont illustrés par les Figures 9.9(b) et 9.9(c) respectivement.

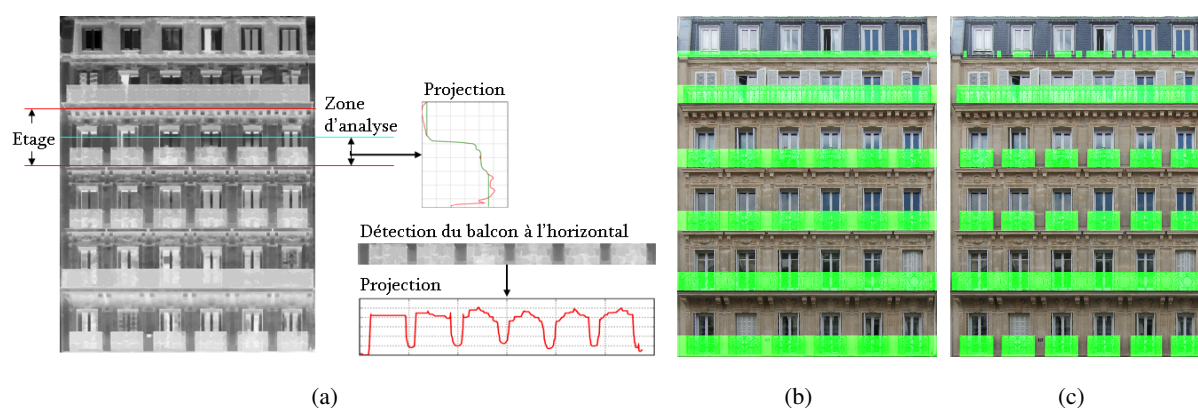


FIGURE 9.9 – (a) Projection horizontale et verticale de l'image. (b) Détection horizontale de balcons et (c) Détection verticale de balcons.

Il faut noter que la méthode présentée dans cette section suppose que les balcons d'un étage sont alignés et ont la même hauteur. Cependant, dans certains bâtiments à cause de la perspective et des différentes profondeurs des balcons, ceux-ci sont décalés dans l'image. C'est pourquoi une étude approfondie par travée est nécessaire.

## 9.4 Modélisation procédurale

Les résultats issus des sections précédentes ont été intégrés dans une plateforme interactive de modélisation procédurale pour la typologie Haussmannienne développée par l'INRIA et THALES dans le cadre du projet TerraNumérica. La plateforme a été implémentée sur Maya en utilisant le graphe de dépendance [77] qui permet une évaluation et une génération dynamique d'un modèle. La modélisation intègre aussi les capacités d'identification et de segmentation des éléments d'architecture d'autres partenaires du projet.

Un exemple de la paramétrisation des règles d'une façade est illustré sur la Figure 9.10. Nous pouvons identifier chacune des étapes de détection présentées : division en étages, détection des dalles, division en travées, détections de balcons et de fenêtres.

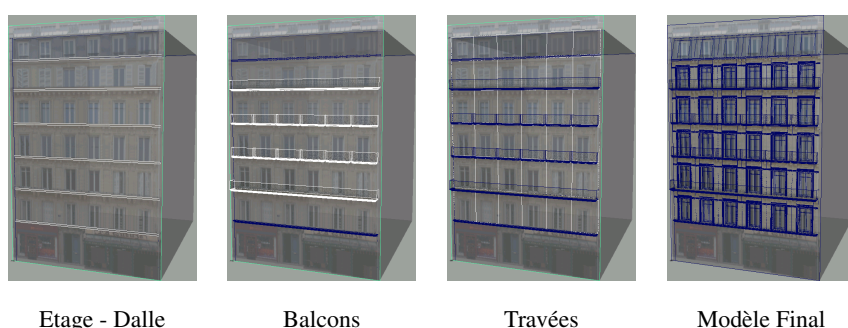


FIGURE 9.10 – Instanciation du modèle de la façade.

Le résultat final du modèle 3D est présenté dans la Figure 9.11. Ces images ont été réalisées par Cédric Guiard de l'INRIA.



FIGURE 9.11 – Rendu 3D du modèle de la façade.

## 9.5 Discussion

Dans ce chapitre nous avons présenté différents outils afin d'extraire l'information sémantique d'une façade : division en étages et en travées, et détection des dalles, des fenêtres et des balcons. Concernant la division par étages, nous avons repris la méthode de division horizontale présentée dans le chapitre précédent et nous l'avons étendue à la détection de la dalle. La méthode détecte une dalle par étage sans tenir compte des possibles décalages produits par la perspective et la profondeur des balcons. C'est pourquoi nous proposons une étape supplémentaire de raffinement. Un premier résultat de cette étape est illustré sur la Figure 9.12. Nous localisons plusieurs dalles sous balcons dans la bonne position ; cependant certains problèmes apparaissent au niveau des ornements et de la partie supérieure de balcons.

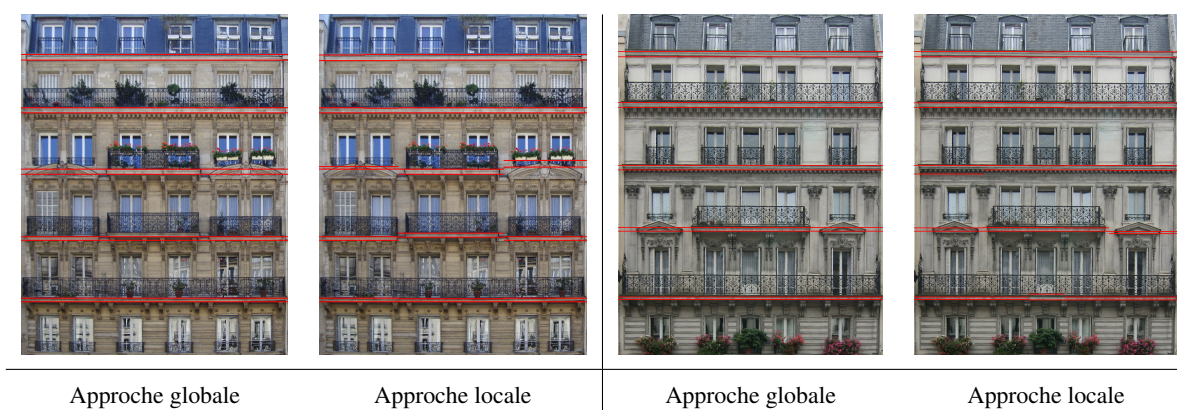


FIGURE 9.12 – Détection de dalles en compensant la profondeur des balcons.

A propos de la division horizontale nous avons assigné des valeurs fixes aux tailles de filtres car nous avons redimensionné chaque image par rapport à sa hauteur avant de lancer la procédure de détection. Concernant la division verticale, la taille du filtre de l'ouverture sur le gradient est fixe pour éliminer les petites structures de bruit. Par contre, la taille du filtre alterné séquentiel *FAS* est automatiquement estimée par rapport à chaque image afin de trouver la division qui minimise l'écart-type de la largeur des travées. De plus, l'utilisation de la série de filtres dans les méthodes rend les algorithmes plus robustes à la texture et/ou petits obstacles. La Figure 9.13 illustre des images de façade avec un obstacle. En utilisant les méthodes proposées dans ce chapitre, nous pouvons remarquer leur robustesse aux petits obstacles. Seulement dans certains endroits de la détection de balcons est erronée, car la toiture et les arbres sont aussi texturés que les balcons. Les résultats sur la base d'images sont présentés dans le site web [cmm.enscm.fr/~hernandez/facade\\_en.html](http://cmm.enscm.fr/~hernandez/facade_en.html).

La méthode pour la détection de fenêtres est réalisée de manière globale sur la façade, en faisant l'hypothèse que les fenêtres sont alignées et qu'il n'y a qu'une fenêtre par travée. Si cela n'est pas le cas, une étape d'analyse et rajustement doit être mise en place.

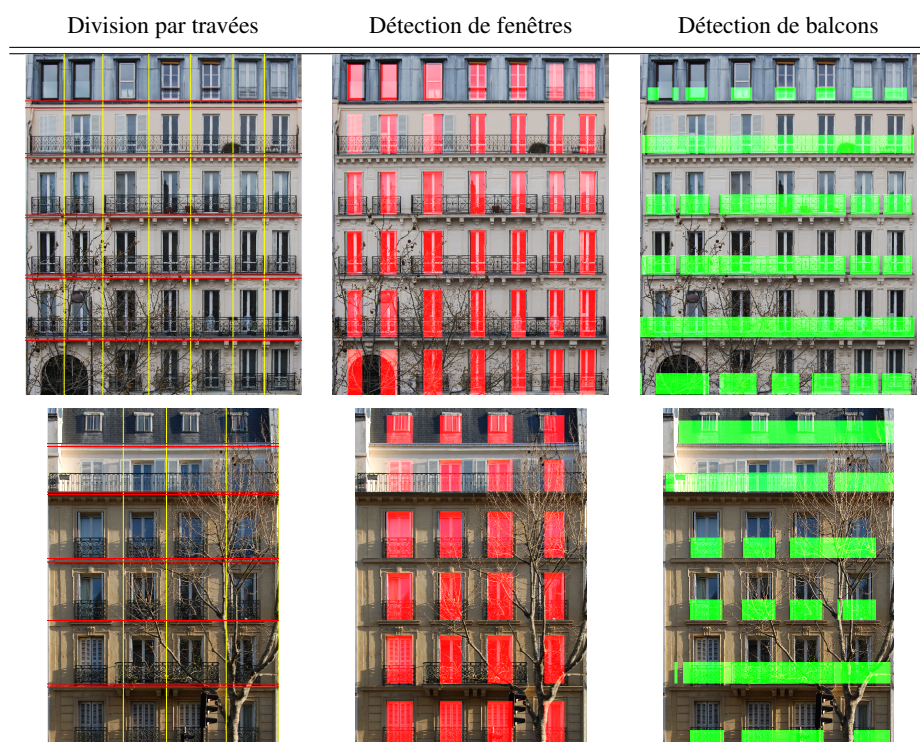


FIGURE 9.13 – Robustesse aux obstacles.

Par ailleurs, nous avons centré la détection de balcons sur les balcons en fer forgé, car ceux-ci sont les plus caractéristiques de l'architecture étudiée. Nous détectons les balcons en profitant du contraste entre eux et des surfaces claires comme les murs en pierre. C'est pour cela que la méthode présente des problèmes pour la détection des balcons sur la toiture. Par ailleurs, nous proposons d'analyser aussi les nuages de points car les balcons ont une différence de profondeur qui peut être exploitée par la suite.

Finalement, les méthodes présentées ont été utilisées comme entrée de la plateforme de modélisation procédurale afin d'extraire à partir des images réelles la sémantique nécessaire pour paramétrer un modèle de façade à l'aide d'une grammaire.

## **Quatrième partie**

# **Analyse des Nuages de Points**

---

# **Segmentation de Nuages de Points**

---

## **Sommaire**

---

<b>10.1 Introduction</b>	<b>138</b>
<b>10.2 Projection sur images</b>	<b>139</b>
<b>10.3 Segmentation par îlots</b>	<b>141</b>
10.3.1 Stratégie N° 1 : Façades alignées	141
10.3.2 Stratégie N° 2 : Façades non-alignées	142
<b>10.4 Segmentation façades-sol</b>	<b>143</b>
<b>10.5 Segmentation de façades en bâtiments</b>	<b>148</b>
<b>10.6 Discussion</b>	<b>149</b>

---

## **10.1 Introduction**

Dans ce chapitre, nous présentons des méthodes pour la segmentation de nuages de points en îlots, sol et façades, et en bâtiments. La segmentation et l'interprétation de ces structures 3D sont parmi les tâches les plus importantes dans la modélisation d'environnements urbains. Par exemple, la modélisation se réalise pour chaque bâtiment, de manière indépendante. Il est donc nécessaire d'avoir les données qui correspondent uniquement aux bâtiments, afin de faciliter son analyse. Les méthodes sont basées sur la projection, pour un point de vue donné, des nuages de points sur des images, où la valeur des pixels correspond en quelque sorte à l'information 3D. Le processus de segmentation est réalisé en 2D, sur l'image que nous avons générée. Une fois la segmentation effectuée, l'image segmentée est rétro-projetée pour découper les nuages de points.

Dans la première section, nous présentons la méthode de projection sur images. La section 10.3 illustre la méthode de segmentation par îlots pour travailler avec des nuages de points denses et gérer les limitations de mémoire. Ensuite, la section 10.4 décrit la méthode qui segmente les données 3D en façades et en sol. Une fois que les données de façades sont séparées des données du sol, il est possible de segmenter les façades par bâtiments comme l'illustre



la section 10.5. Dans la dernière section, nous discutons plus en détail les caractéristiques de nos approches.

## 10.2 Projection sur images

Comme nous l'avons déjà mentionné, les méthodes développées dans ce chapitre s'appuient sur la projection de nuages de points sur des images. Ces images sont générées en utilisant un modèle de caméra "virtuelle". Le modèle de caméra  $\mathcal{P}$  est une transformation "projective" de  $\mathbf{R}^3 \rightarrow \mathbf{N}^2$ . Cette transformation peut être décomposée en trois transformations successives de repère comme l'illustre la Définition 18.

**Définition 18.** Soit  $M = \langle X, Y, Z \rangle$  un point 3D de l'espace  $\mathbf{R}^3$  et  $m = \langle u, v \rangle$  un point-image de l'espace  $\mathbf{N}^2$ , la fonction  $\mathcal{P}$  est définie par :

$$\langle X, Y, Z \rangle \xrightarrow{T} \langle X_c, Y_c, Z_c \rangle \xrightarrow{P} \langle x, y \rangle \xrightarrow{A} \langle u, v \rangle \quad (10.1)$$

$$[T] = \begin{bmatrix} [R] & t \\ 0^T & 1 \end{bmatrix}$$

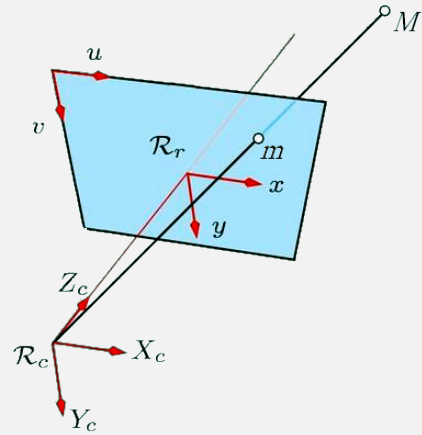
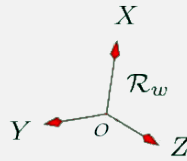
$R$  rotation et  $t$  translation.

$$[P] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$

$f$  focale de l'objectif.

$$[A] = \begin{bmatrix} k_x & 0 & c_x \\ 0 & k_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

$c_x, c_y$  (en pixels) coordonnées de l'intersection de l'axe optique avec le plan image.  $k_x, k_y$  nombre de pixels par unité de longueur.



où,  $\mathcal{R}_w$  est le repère de l'espace réel,  $\mathcal{R}_c$  est le repère de la caméra (dont l'origine est située au centre optique de la caméra) et  $\mathcal{R}_r$  est le repère du capteur (plan de la rétine).

L'analyse principale 3D est faite à partir d'une vue supérieure de la scène, c'est-à-dire que la caméra "virtuelle" choisie a les caractéristiques suivantes :

- L'axe optique de la caméra est parallèle à l'axe  $z$  de la scène et le plan de la caméra passe par le point  $(0, 0, z_{min})$ . (La matrice de rotation  $[R]$  est égale à l'identité).
- Le nuage de points est centré pour faire coïncider son centre de gravité avec le centre de l'image. (Le vecteur de translation  $t$  est égal au centre de gravité des points et  $c_x = c_y = 0$ ).
- La projection est parallèle orthographique : l'axe de projection est orthogonal au plan de projection et le centre de projection se trouve à l'infini. ( $f = \infty$  et donc  $x = X_c, y = Y_c$ ).

Afin d'éviter les problèmes d'échantillonnage, les valeurs  $k_x$  et  $k_y$  doivent être soigneusement choisies. Si elles sont trop petites, plusieurs points seront projetés sur les mêmes coordonnées de l'image et il en résultera une importante perte d'information. Par contre, si elles sont trop larges, la connexité de pixels, requise par nos méthodes, n'est pas assurée. Par conséquent, le choix idéal des dimensions est de 1 : 1 ; autrement dit, un pixel pour chaque point 3D dans le plan de la caméra. Après avoir choisi la projection, plusieurs mesures peuvent être assignées à la valeur d'intensité du pixel dans l'image. Cette valeur est fonction des points 3D projetés sur les mêmes coordonnées  $u, v$ . Par exemple la distance maximale des points au plan de la caméra [44] (notion de profondeur), la différence entre le maximum et le minimum (notion de hauteur) ou le nombre de points (notion d'accumulation).

La Figure 10.1 montre un exemple de la projection d'un nuage de points sur plusieurs images : profondeur, hauteur et accumulation. Dans ce cas, le nuage de points a une résolution moyenne d'environ  $5cm$  et pour cette raison, la valeur que nous avons sélectionnée est  $k_x = k_y = 20pix/m$ . Notez que dans les images, les valeurs maximales se trouvent dans la région des façades. Ce comportement corrobore le fait que les façades sont les structures les plus hautes et où la plupart des points sont projetés.

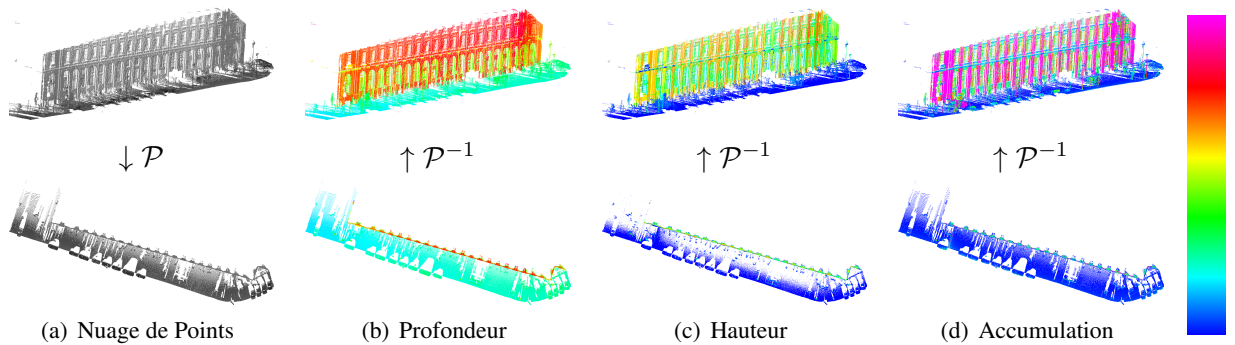


FIGURE 10.1 – (a) Nuages de points - vue 3D et vue supérieure. Projection sur images (code couleurs au rang) : (b) la valeur maximale – rang (0, 23651mm), (c) la différence max – min – rang (0, 18816mm) et (d) l'accumulation – rang (0, 121fois). Données ©IGN Tronçon de la Rue Soufflot côté pair.

L'exemple présenté est un tronçon de rue de dimensions XY égales à  $103,5m \times 48,4m$ , et donc les images produites ont une taille de  $2070pix \times 968pix$ . Dans le cas réel, les systèmes d'acquisition fournissent des nuages de points de plusieurs centaines de mètres. Néanmoins, lorsque nous essayons de produire les images de ces nuages de points, nous pouvons dépasser les limites de mémoire disponible<sup>1</sup>. En conséquence, une étape de découpage préalable à partir des images de profondeur de basse résolution est mise en place. Le découpage doit être cohérent avec les structures d'intérêt comme les bâtiments et c'est pour cela que nous nous intéressons à la séparation par

1. Nous pouvons créer des images de dimensions d'indexation jusqu'à 26 bits  $((2^{32} - 1) / (\text{sizeof}(\text{double})))$  pour le cas d'un ordinateur 32 bits



pâtés de maisons.

### 10.3 Segmentation par îlots

Pour le découpage en pâtés de maisons, on peut exploiter l'information des images avec une résolution réduite de 4 : 1 de manière à éviter les problèmes de mémoire. La méthode est illustrée dans le diagramme de la Figure 10.2. D'abord, le nuage de points est projeté avec une caméra  $\mathcal{P}$  de la vue supérieure de la scène. On a choisi une vue supérieure car dans cette image les îlots ne sont pas connectés entre eux. Puis, nous extrayons et analysons les informations de façades qui caractérisent les îlots, afin d'appliquer la segmentation. Pour ces étapes d'analyse et de segmentation nous avons développé deux stratégies qui seront expliquées postérieurement. A partir de l'image segmentée, les régions sont rétro-projetées sur le nuage de points avec la fonction inverse de caméra  $\mathcal{P}^{-1}$ .

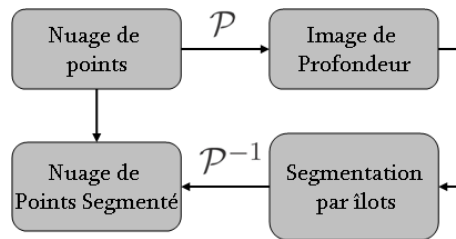


FIGURE 10.2 – Les étapes de la segmentation par îlots.

Deux stratégies ont été mises en place pour l'analyse de l'information d'une façade qui dépend de l'alignement global des façades des données à traiter. Les deux cas d'alignement sont illustrés sur la Figure 10.3. Dans le premier cas, nous supposons que les données correspondent à un seul côté de rue et les différentes façades sont alignées. Ce cas sera illustré dans la section 10.3.1. Dans la deuxième stratégie, la condition d'alignement n'est pas nécessaire, néanmoins plusieurs contraintes de taille (largeur et hauteur) sont utilisées (voir section 10.3.2).

#### 10.3.1 Stratégie N° 1 : Façades alignées

Cette méthode est appliquée sur des nuages de points d'un seul côté de la rue où les façades sont alignées. Nous générons l'image de profondeur, en assignant à chaque pixel de l'image la distance maximale entre les points 3D projetés sur le même pixel et le plan de caméra. L'algorithme utilise la transformée de Hough pondérée par la valeur d'intensité de chaque pixel. Cette procédure nous permet de détecter la ligne de direction des façades sur l'image de profondeur (voir Figure 10.4(a)). Sur cette ligne, nous analysons le profil des profondeurs des bâtiments. Nous détectons les variations de profondeur importantes, qui correspondent aux rues séparant les différents pâtés de maisons. Dans cette analyse, le profil est d'abord filtré, en utilisant un filtre alterné séquentiel, pour réduire le bruit et faciliter la détection. La Figure 10.4(b) montre le profil original et le profil filtré. Ensuite, le profil filtré est

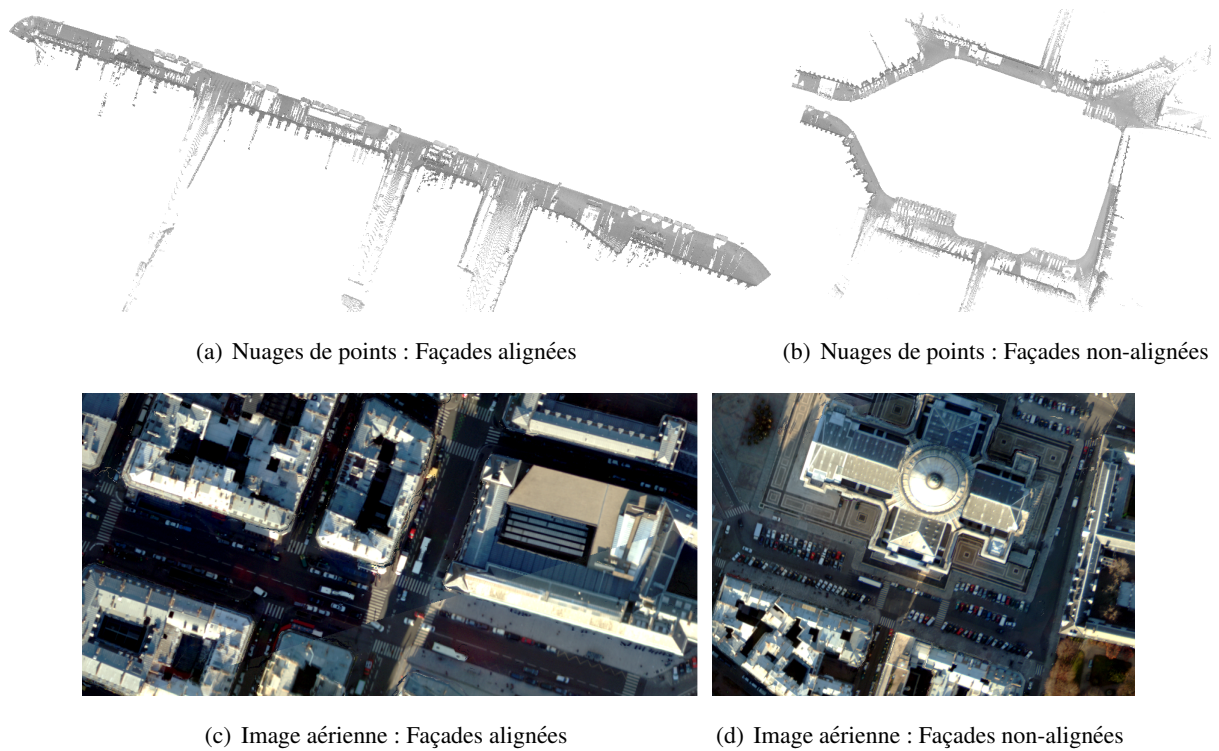


FIGURE 10.3 – Alignement global de façades. (a)- (c) rue Soufflot, (b)- (d) place du Panthéon.

segmenté en utilisant la ligne de partage des eaux. Ces divisions sont tracées dans la direction perpendiculaire à la direction de la façade pour découper l'ensemble en tronçons de rue. L'image segmentée est ensuite rétro-projetée pour découper le nuage de points 3D original (voir Figure 10.4(c)).

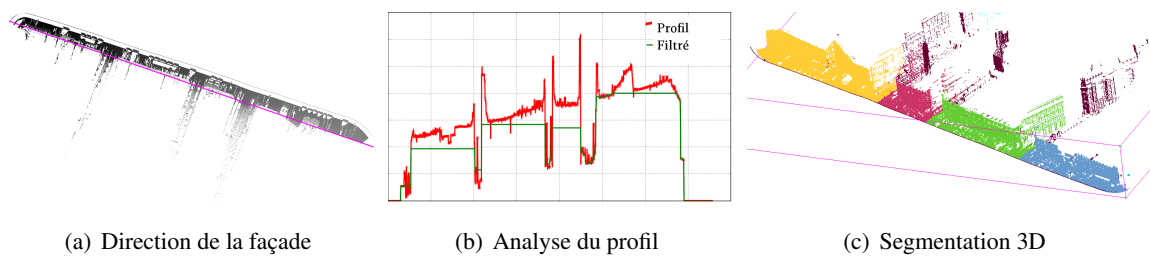


FIGURE 10.4 – Images intermédiaires de la segmentation par îlots de rues alignées.

### 10.3.2 Stratégie N° 2 : Façades non-alignées

Dans cette méthode, nous faisons plusieurs hypothèses : la hauteur absolue d'un bâtiment est d'au moins  $4m$ , la distance minimale entre îlots est supérieure à  $2m$  et la longueur d'un îlot est d'au moins  $10m$ . La procédure est la

suivante : d'abord, nous générons l'image projetée, en assignant à chaque pixel la différence des distances maximum et minimum des points projetés sur les mêmes coordonnées de l'image. Ceci nous donne une information de la hauteur des structures urbaines sans que nous soyons perturbés par les rues en pente. Ensuite, nous appliquons un seuil de la hauteur minimale d'un bâtiment. L'image résultat présente des *CC*s de bâtiments et des objets hauts tels que les lampadaires ; néanmoins dans un même îlot, il peut y avoir de petites déconnexions entre façades, dues à des données manquantes. Pour fusionner les façades dans les îlots, nous appliquons une fermeture d'une taille de  $2m$ . Puis nous éliminons les *CC*s d'élongation inférieure à  $10m$ . Les *CC*s résultantes correspondent aux îlots de la scène. Pour tracer la séparation de manière équidistante entre îlots voisins, nous calculons la fonction distance aux *CC*s détectées et lui appliquons la ligne de partage des eaux. L'image segmentée est également rétro-projetée pour découper le nuage de points original.

Les Figures 10.5 et 10.6 illustrent la segmentation par îlots de données 3D issues des deux systèmes d'acquisition, Lara3D et Stereopolis respectivement. Deux cas sont illustrés : deux côtés de rues et plusieurs îlots en diverses directions. Comme l'on peut observer, la méthode présente de bons résultats dans les deux cas. Néanmoins, dans le cas de la Figure 10.5 correspondant aux données CAOR, on a détecté séparément les régions 1 et 2 qui appartiennent au même îlot (de même pour les régions 5, 6 et 7), car le profil d'acquisition du capteur laser est perpendiculaire à la direction de la façade et il y a trop de données manquantes. Ces régions sont séparées par une distance supérieure à  $2m$  et la fermeture utilisée n'arrive pas à les relier.

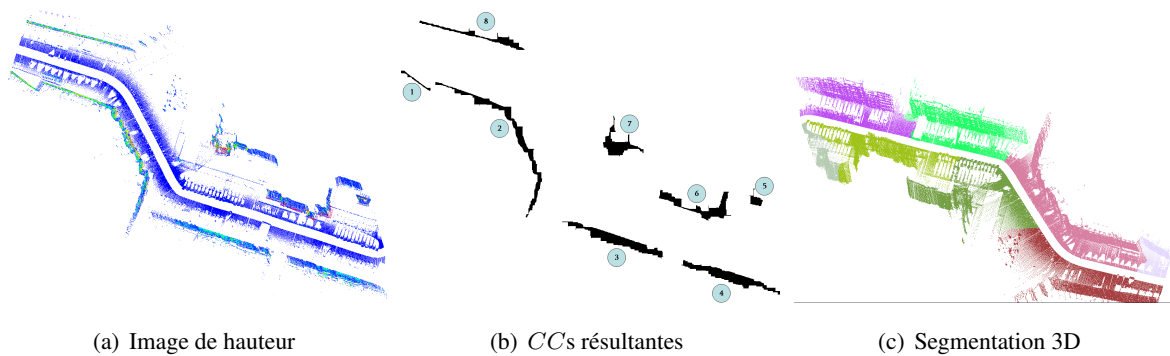


FIGURE 10.5 – Segmentation par îlots de rues non-alignées. Données ©CAOR Fin de la Rue Soufflot – début de la Place du Panthéon.

## 10.4 Segmentation façades-sol

Une autre étape également intéressante à analyser est la segmentation de points 3D entre la façade et le sol (rue et trottoir). Cette segmentation est importante parce qu'elle facilite plusieurs étapes de la modélisation urbaine :

- La modélisation de la façade.

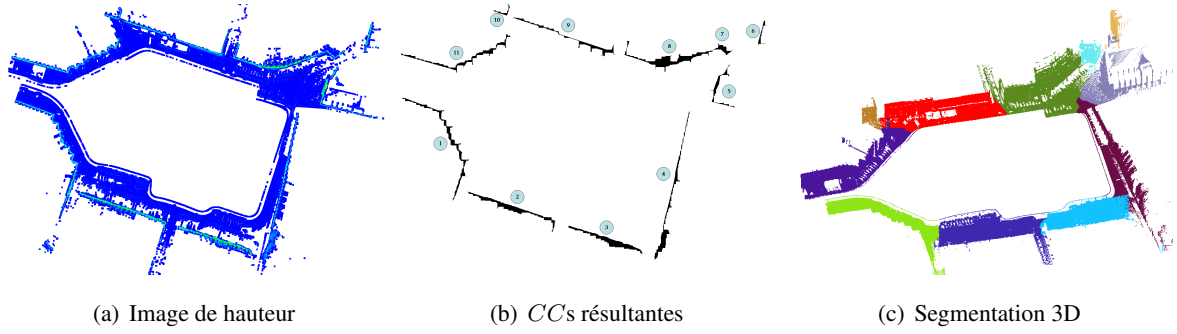


FIGURE 10.6 – Segmentation par îlots de rues non-alignées. Données ©IGN Place du Panthéon.

- L'extraction de la façade pour chaque bâtiment à partir d'une transformation de caméra (voir section 10.5).
- La détection d'artefacts au niveau du sol (voir Chapitre 11).
- La segmentation route-trottoir et la modélisation de la chaussée (voir Chapitre 12).

Nous travaillons avec les données déjà découpées en pâtes de maison et avec une résolution idéale d'image  $20\text{pix}/\text{cm}$ . Néanmoins les méthodes présentées pourront être appliquées aux données brutes si la capacité de mémoire disponible est suffisante.

Cette approche est basée sur une segmentation grossière de l'image de profondeur, réalisée au moyen d'un algorithme de zones quasi-plates ( $\lambda$ - zones plates<sup>2</sup>) présenté par Meyer dans [107].

**Définition 19.** Deux pixels voisins  $p$  et  $q$  appartiennent à la même zone quasi-plate d'une image  $I$ , si leur différence de profondeur  $|I_p - I_q|$  est inférieure à une distance  $\lambda$  donnée.

$$\forall (p, q) \text{ voisins} : |I_p - I_q| \leq \lambda \quad (10.2)$$

Etant donné que les variations de profondeur au niveau du sol sont faibles, la segmentation de l'image de profondeur avec une distance  $\lambda \approx 1\text{m}$  réunira les pixels du sol en une seule région, en incluant les artefacts qui s'y trouvent. Cependant, à cause des occlusions et des caractéristiques des systèmes d'acquisition, il y a des informations manquantes au niveau du sol. Ces données manquantes limitent la performance de la méthode car la connexité des pixels du sol n'est pas assurée. De manière générale, la taille des informations manquantes dépend principalement de la hauteur de l'obstacle, de la distance du capteur au sol et de la distance du capteur à l'obstacle, comme la Figure 10.7 l'illustre.

L'image présente le processus d'acquisition par des systèmes mobiles d'un trottoir en pente et avec deux obstacles. Dans cet exemple, le sol est divisé en une région de rue par la voiture et deux régions de trottoir par le piéton. La figure illustre également un profil obtenu par la projection du point 3D le plus haut (profil de profondeur de la vue

2. L'idée de base de la méthode a été introduite par Nagao *et al.* dans [115]. D'ailleurs, Soille dans [152] la décrit comme  $\alpha$ -composantes connexes.

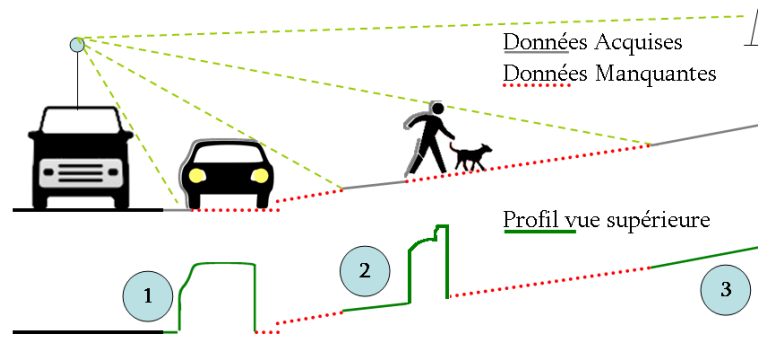


FIGURE 10.7 – Illustration de l'acquisition

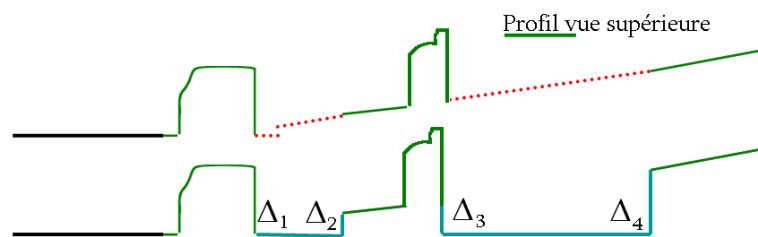


FIGURE 10.8 – Assignment des pixels manquants avec la valeur minima.

supérieure).

Ce problème se présente avec les données réelles comme l'illustre la Figure 10.10(a). La figure montre les images de profondeur de la vue supérieure des trois systèmes d'acquisition. Par exemple, on peut observer dans l'illustration que l'information manquante est plus importante lorsque l'objet est éloigné du capteur. D'ailleurs, les données issues du système VX ©MENSI sont les données qui présentent le moins de points du sol parce que le capteur fixe a été dirigé vers la façade. De plus, ces données ont une résolution plus faible que les deux autres systèmes.

Une première solution serait d'assigner les pixels non-acquis avec la valeur minimale de l'image de profondeur. Néanmoins, cette solution ne résout pas le problème de manière générale. Analysons le cas du trottoir incliné avec l'algorithme de  $\lambda$ - zones plates (voir Figure. 10.8). Une valeur  $\lambda < \Delta_3$  ne réunirait pas toutes les régions du sol. Par contre, une valeur de  $\lambda \geq \Delta_3$  trop grande fusionnerait les régions du sol, mais elle pourrait aussi fusionner les points de la façade. Dans les données sur lesquelles nous travaillons, une segmentation de régions quasi-plates, avec une valeur  $\lambda > 1m$ , présente des fuites vers les façades. C'est pour cela que le remplissage des pixels non-acquis doit être appliqué de manière locale, par rapport aux régions voisines.

En morphologie mathématique, il existe une technique qui permet de boucher des régions dites «trous». Un trou est un ensemble de pixels dont le minimum n'est pas connecté au bord de l'image  $I$  [151]. L'algorithme de remplissage

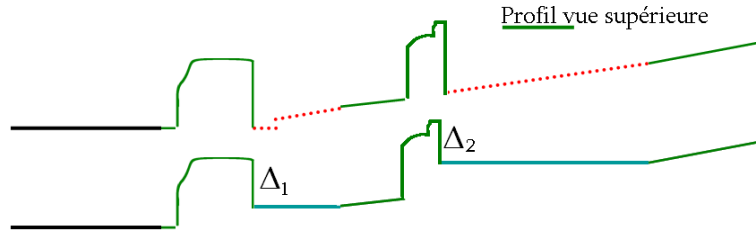


FIGURE 10.9 – Assignment des pixels manquants avec le remplissage de trous.

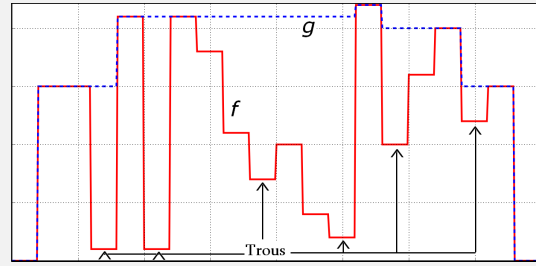
de trous consiste donc à supprimer les minima qui ne sont pas liés au bord de l'image. Il utilise une reconstruction morphologique par érosion ( $R_f^e$ ) où le marqueur est une image à la valeur maximale partout sauf sur son bord (voir Définition 20).

**Définition 20.** Soit  $f$  une fonction et  $p$  un point quelconque dans  $f$ . Le remplissage de trous  $g$  présenté par Soille [150] est défini par :

$$g = \text{Fill}(f) = R_f^e(m_k) \quad (10.3)$$

où,

$$m_k = \begin{cases} f_p & \text{Si } p \text{ est au bord} \\ \max(f) & \text{autrement} \end{cases} \quad (10.4)$$



Observons le résultat dans le cas du trottoir incliné avec le remplissage de trous comme l'illustre la Figure 10.9. Dans ce cas, la valeur de  $\lambda$  nécessaire pour fusionner les régions du sol, y compris les objets, est inférieure à la première solution présentée antérieurement. C'est pourquoi une valeur de  $\lambda = 1m$  sur les données réelles peut être choisie afin d'éviter les fuites vers la façade.

Ainsi, en utilisant l'algorithme de remplissage, on peut donc boucher les trous qui ne touchent pas le bord (voir Figure 10.10(b)). Malgré cela, comme on peut le constater dans la Figure 10.10(b), seuls quelques trous des données STEREOPOLIS ont été remplis. Pour mieux comprendre ce problème voici une description des différents types de données :

- Dans le cas des données MENSİ, les données du sol sont de petites régions isolées et éparpillées.
- Dans le cas des données de systèmes mobiles, pour une scène comprenant beaucoup d'obstacles (par exemple des voitures garées) ou de grands obstacles (des lampadaires), les trous sont connectés au bord de l'image.

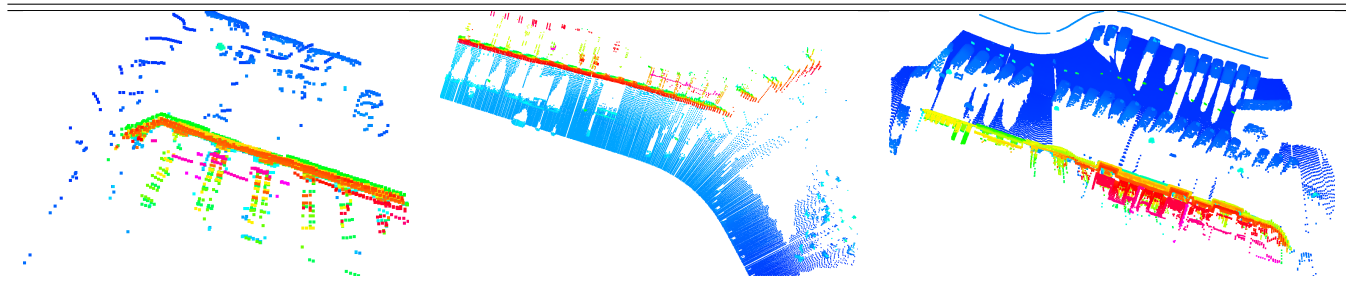
Pour résoudre ce problème, nous proposons de relier les régions du sol en utilisant le chemin de la distance minimale entre elles. Autrement dit, une région sera reliée à sa voisine la plus proche. Le résultat final des connexions

Systèmes d'acquisition

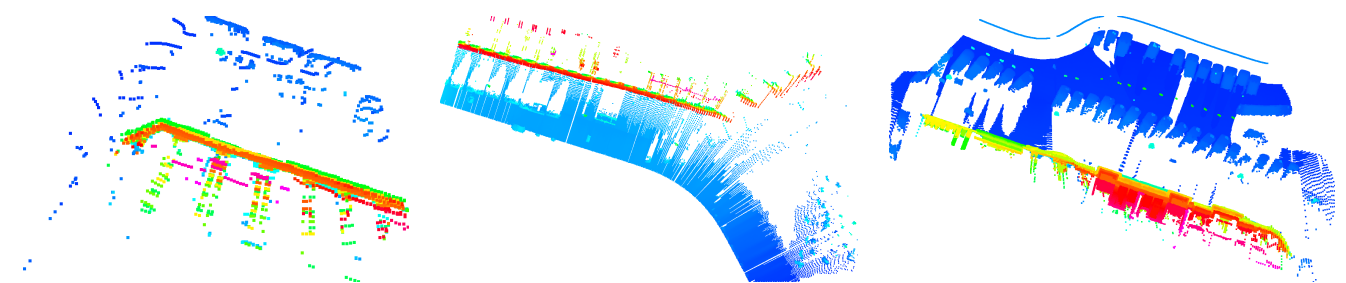
VX ©MENSI

Lara3D ©CAOR

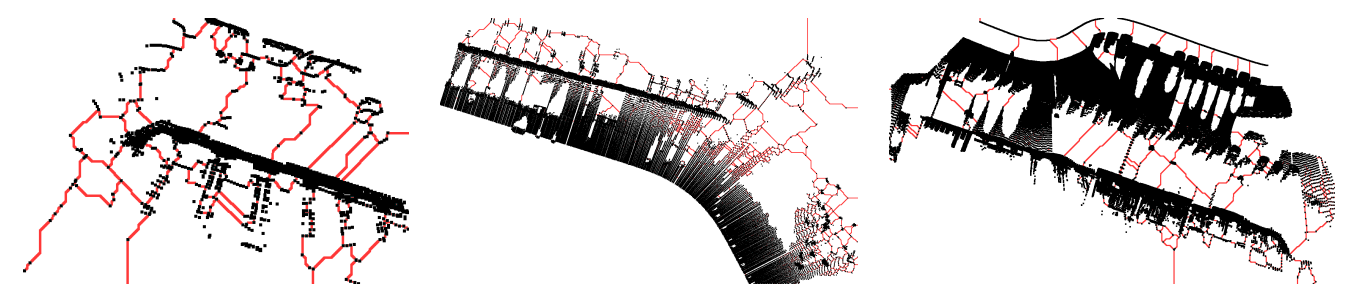
Stereopolis ©IGN



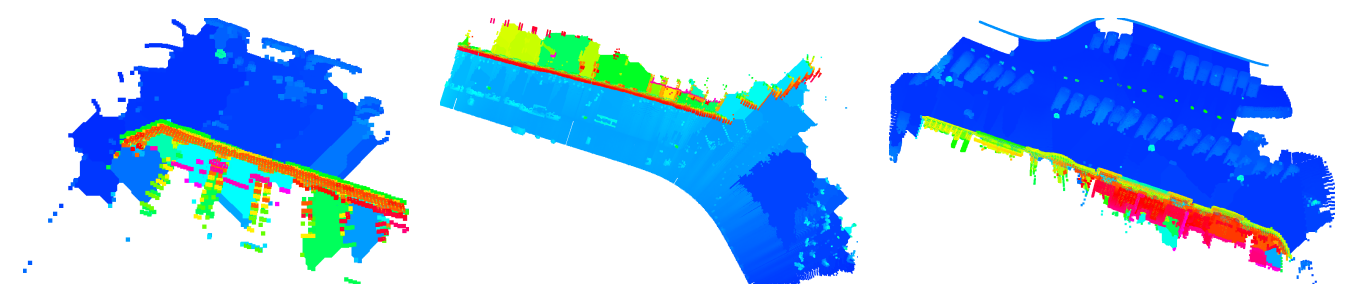
(a) Images de profondeur



(b) Remplissage de trous



(c) Connexion de régions



(d) Remplissage de trous après la connexion de régions

FIGURE 10.10 – Estimation des pixels manquants en utilisant le remplissage de trous.



## Systèmes d'acquisition

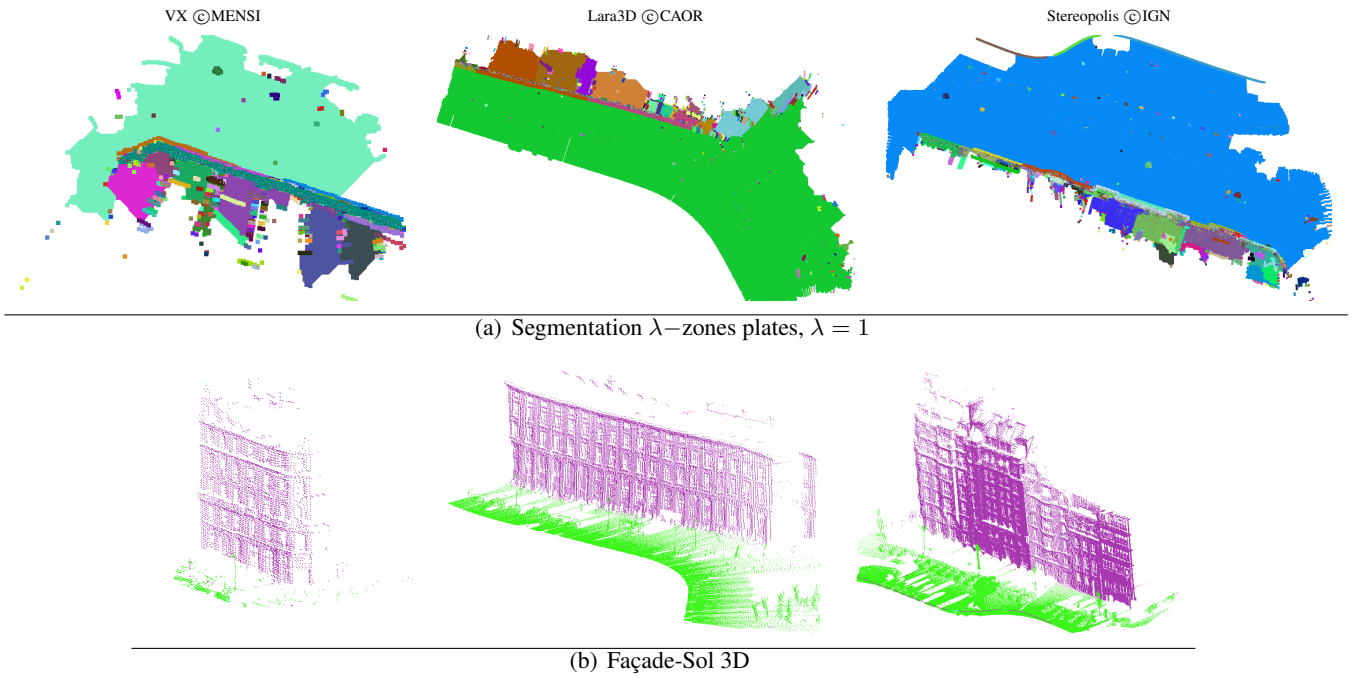


FIGURE 10.11 – Segmentation Façade-Sol.

est illustré dans la Figure 10.10(c). Maintenant, les régions et leurs connexions forment de nouveaux trous qui doivent être bouchés. Le remplissage s'applique bassin par bassin de manière indépendante afin de réaliser une estimation locale de la région à combler. La Figure 10.10(d) montre le nouveau remplissage de l'image de profondeur.

En utilisant l'image comblée, nous pouvons appliquer la segmentation de zones quasi-plates (voir Définition 10.2), en obtenant comme résultat les images de la Figure 10.11(a). En choisissant la région la plus grande, nous obtenons le masque du sol. Cette première région présente quelques trous produits par des éléments très hauts comme les lampadaires. Afin de récupérer ces éléments, on assigne au masque du sol tous ceux qui sont contenus dans le premier masque. Le masque final est rétro-projeté sur le nuage de points, permettant la segmentation de nuages de points 3D de façades et points du sol (voir Figure 10.11(b)).

## 10.5 Segmentation de façades en bâtiments

Nous avons déjà décrit dans le Chapitre 8 une technique automatique pour séparer/segmenter la façade d'un bâtiment à partir d'une image. D'ailleurs, comme nous l'avons déjà mentionné, les images de texture sont géo-référencées et recalées aux nuages de points avec une matrice de caméra  $\mathcal{P}$ . En projetant la séparation des façades,



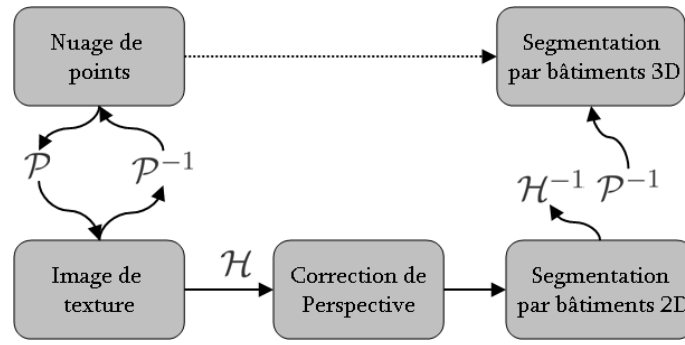


FIGURE 10.12 – Les étapes de la segmentation de façades en bâtiments.

on pourra donc séparer les données 3D. Rappelons que le découpage dans l'image est fait à partir d'une image rectifiée. Cette correction de perspective est réalisée en utilisant une matrice d'homographie  $\mathcal{H}$ . Lorsque nous voulons projeter la segmentation, les séparations au niveau de l'image rectifiée doivent passer au repère de l'image originale avec la transformation  $\mathcal{H}^{-1}$ . Cette procédure est décrite par le diagramme de la Figure 10.12.

Un exemple de cette segmentation est illustré sur la Figure 10.13. Les données de façade de la Figure 10.13(b) sont la sortie de la méthode de segmentation façades-sol (voir section 10.4). Cette segmentation préalable est nécessaire afin d'éliminer les structures 3D qui se trouvent devant la façade par rapport au point de vue de la caméra  $P$ . Pour illustrer ce phénomène, on a coloré le nuage de points avec l'image de texture et en rouge les points qui ne sont pas dans l'image. Le résultat final de segmentation par bâtiment est présenté dans la Figure 10.13(d).

Par ailleurs, toute information disponible sur l'image peut aussi être rétro-projetée sur les données 3D en utilisant la même technique décrite par le diagramme de la Figure 10.12. Un exemple est la projection des séparations horizontales par étage. La Figure 10.14(a) montre les divisions horizontales de la façade. Ces divisions sont projetées comme l'illustre la Figure 10.14(c).

## 10.6 Discussion

Dans ce chapitre, nous avons décrit trois méthodes de segmentation de nuages de points en îlots, chaque îlot en points façade et en points sol ; chaque façade d'un îlot en bâtiments isolés. Eventuellement, nous pourrions traiter un nuage de points sans la segmentation préalable en îlots, à condition de ne pas dépasser la quantité de mémoire disponible. Les méthodes sont basées sur la projection sur images, principalement en utilisant une vue supérieure comme caméra «virtuelle». Les méthodes sont complètement automatiques avec les mêmes paramètres. C'est pourquoi l'information des façades doit se trouver en grande partie dans l'axe  $z$  et les données du sol doivent

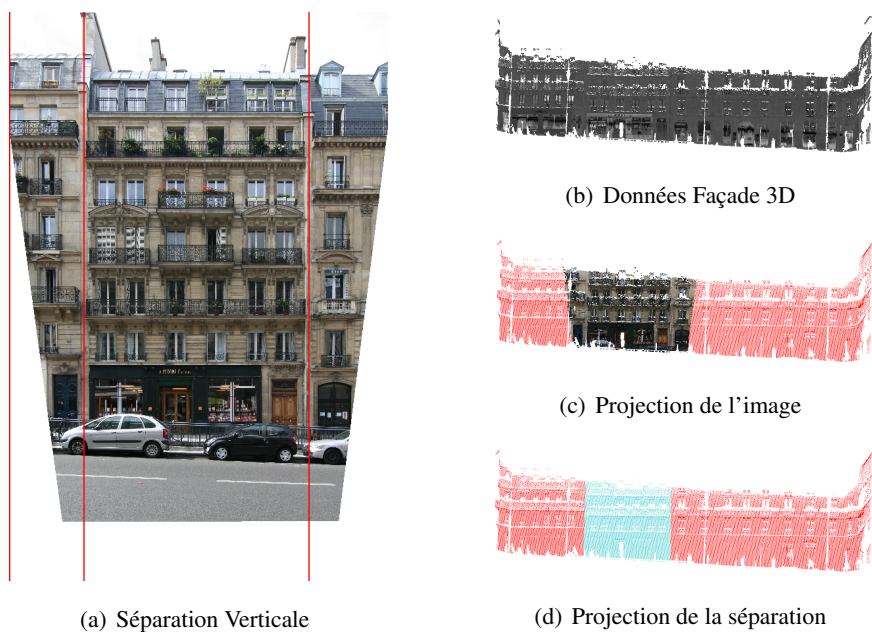


FIGURE 10.13 – Segmentation de façades en bâtiments. Données 3D ©IGN, Image 2D ©MENSI Rue Soufflot - Bâtiment N-13.

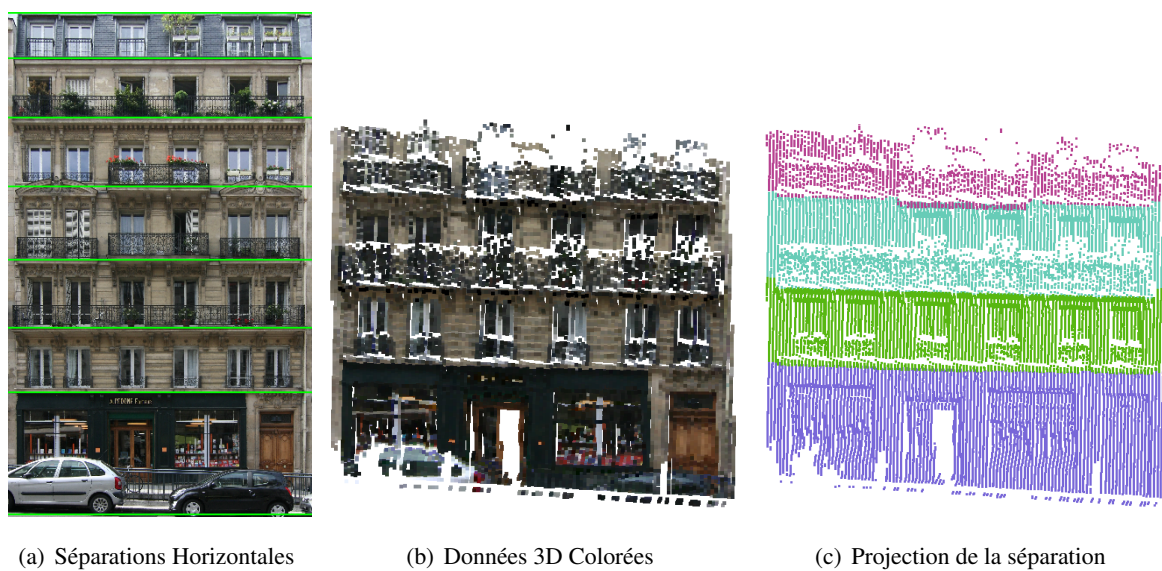


FIGURE 10.14 – Segmentation des étages de façades.

être perpendiculaires aux façades. C'est normalement le cas pour les systèmes d'acquisition de données urbaines géo-référencées. C'est pour cela qu'une des qualités des méthodes présentées est l'indépendance aux systèmes d'acquisition, par exemple on obtient de bons résultats avec trois systèmes de caractéristiques différentes.

Pour le cas des données issues de systèmes mobiles, la segmentation par îlots sépare les nuages de points en pâtés de maisons afin de faciliter l'analyse de données 3D denses. Deux stratégies ont été mises en place par rapport à l'alignement des façades dans les données à traiter. La première stratégie est non-paramétrique ; néanmoins, la technique est appliquée uniquement sur un seul côté de rue et avec une contrainte d'alignement de façades. La deuxième stratégie utilise des paramètres caractéristiques des scènes urbaines tels que la hauteur minimale d'un bâtiment, la longueur minimale d'un îlot, et la distance minimale qui les sépare. Cette technique peut être appliquée sur toutes les variétés de données 3D issues de systèmes utilisés présentés dans cette thèse. La principale faiblesse de ces stratégies est l'information manquante des façades, nécessaire pour assurer la connexion dans un îlot donné. Ce problème est particulier aux bâtiments de formes complexes comme les monuments historiques.

La segmentation façades-sol, comme son nom l'indique, cherche à séparer les données façades de données sol, y compris les artefacts. La méthode utilise principalement deux opérateurs morphologiques, le remplissage de trous et la segmentation de zones quasi-plates. Le remplissage de trous est utilisé après une connexion de régions isolées, afin d'estimer la connexité du sol de manière locale. La segmentation de  $\lambda$  – zones plates utilise une valeur  $\lambda = 1m$  pour garantir la fusion de pixels du sol, sans présenter de fuites vers les façades.

La méthode de segmentation d'un bâtiment utilise l'information de caméra du géo-référencement et la segmentation de l'image 2D décrite au Chapitre 8. La méthode est appliquée aux données de façades sans les artefacts qui se trouvent dans la même ligne de projection de la caméra. Dans cette méthode, aucun traitement additionnel au niveau de la segmentation n'est effectué, excepté la rétroprojection de la segmentation 2D. C'est pourquoi les erreurs produites dans le découpage de la façade seront propagées au nuage de points. De plus, si l'image de la caméra n'est pas soigneusement calibrée, le découpage 3D produira des résultats erronés. Une voie d'exploration possible serait une sorte de segmentation coopérative, c'est-à-dire qu'elle utilise à la fois l'information 2D et 3D afin de rendre plus robuste la segmentation.

# Détection et Classification d'Artefacts

## Sommaire

<b>11.1 Introduction</b>	<b>152</b>
<b>11.2 Détection et filtrage</b>	<b>153</b>
<b>11.3 Classification d'artefacts</b>	<b>156</b>
11.3.1 Artefacts connectés	156
11.3.2 Analyse de caractéristiques	156
11.3.3 Classification	157
<b>11.4 Discussion</b>	<b>159</b>
11.4.1 Artefacts cachés et non-séparation d'artefacts	162
11.4.2 Non-séparation d'artefacts	162

## 11.1 Introduction

Dans le chapitre précédent, nous avons présenté des techniques de segmentation de nuages de points en utilisant la projection sur des images. Ce chapitre illustre une méthode automatique de détection et d'interprétation d'artefacts au niveau du sol, comme les véhicules, les piétons ou le mobilier urbain. Comme nous l'avons déjà mentionné dans l'état de l'art, section 2.3.2, les approches existantes quant à la détection et au filtrage d'artefacts présentent de fortes contraintes selon le système d'acquisition [56, 97] ou se limitent à un seul type d'objet à détecter dans la scène [30, 121, 175]. Dans notre cas, l'acquisition de données du projet est réalisée sans intervenir sur l'activité naturelle de la ville et avec trois systèmes d'acquisition de caractéristiques différentes.

La détection des artefacts a plusieurs objectifs :

- Le filtrage de ces structures pour faciliter l'étape de modélisation de bâtiments, de façades et de sol.
- L'utilisation des éléments fixes, tels que les lampadaires, comme points de repère dans une procédure de recalage entre plusieurs systèmes ou entre plusieurs acquisitions d'un même système.
- La réintroduction de certains éléments du mobilier urbain, améliorant le réalisme visuel de la scène modélisée.

La méthode de détection des artefacts est réalisée sur l'image de profondeur, néanmoins d'autres informations comme l'image d'accumulation et l'image de hauteur seront utilisées pour l'extraction de caractéristiques supplémentaires. La séparation façades-sol présentée dans le chapitre précédent, sera prise en compte dans ce chapitre. Nous faisons en effet l'hypothèse que les artefacts sont positionnés sur le sol.

Dans la première section, nous décrivons la méthode de détection d'artefacts. La section 11.2 illustre dans un second temps la méthode de classification. La dernière section détaille les caractéristiques de nos approches.

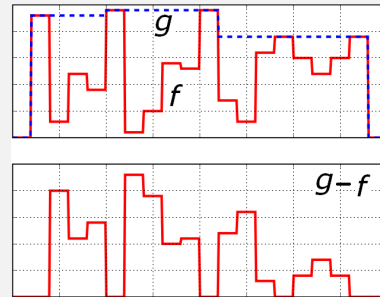
## 11.2 Détection et filtrage

Nous faisons l'hypothèse que les artefacts se trouvent au sol. C'est pourquoi nous utilisons la séparation façade-sol pour extraire un masque du sol. Ensuite, nous ne traitons que les artefacts à l'intérieur de ce masque. Comme image de départ nous utilisons l'image de profondeur comblée que nous avons utilisée pour la détection du sol (voir la Figure 10.10(d)). Avec cette étape, l'image de profondeur est plus homogène car cette transformation réduit également le bruit sur les artefacts.

La méthode de détection est basée sur le chapeau haut de forme de l'algorithme de remplissage de trous (voir Définition 21).

**Définition 21.** Soit  $f$  une fonction et  $p$  un point quelconque dans  $f$ . Le chapeau haut de forme du remplissage de trous  $\text{FTH}(f)$  est défini par :

$$\text{FTH}(f) = \text{Fill}(f) - f \quad (11.1)$$



Analysons l'illustration du trottoir incliné de la Figure 10.9. Les deux artefacts sont des bosses sur le profil. Ainsi, si nous inversons le profil, ces bosses deviennent des trous. En utilisant l'opérateur de chapeau haut de forme, après un remplissage de trous du profil inversé, nous obtenons une détection d'artefacts avec une estimation de leur hauteur. Comme on peut l'observer sur l'illustration de la Figure 11.1, les deux artefacts sont correctement détectés.

Même si on a décrit la méthode sur un profil, cette technique est utilisée en 2D sur l'image de profondeur remplie. Néanmoins, deux détails additionnels doivent être considérés :

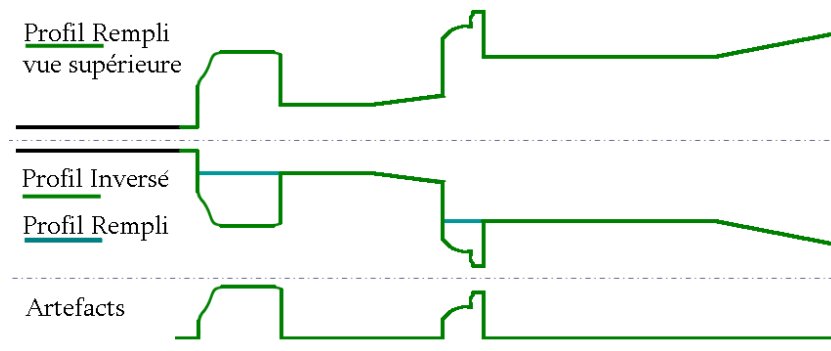


FIGURE 11.1 – Détection d'artefacts du profil 10.9.

- Une fois que l'image est inversée, le fond de l'image de profondeur (noir) devient la valeur maximale. De cette manière, toute l'information de profondeur est un grand trou. Pour résoudre ce problème nous inversons seulement les points différents de zéro (les pixels à zéro restent à zéro).
- Les artefacts qui se trouvent à la frontière, spécialement les voitures, manquent d'information latérale à cause de la proximité des objets au capteur. En appliquant le chapeau haut de forme, l'artefact n'aura que l'information de la partie haute. Le problème est corrigé en ajoutant, autour du masque du sol du côté de la rue, un bord à la hauteur minimale de l'image de profondeur.

Un exemple qui illustre ces problèmes et leurs solutions est présenté dans la Figure 11.2. Dans l'illustration deux artefacts sont présents et l'un d'eux est trop proche du système d'acquisition. Comme nous l'avons mentionné, si nous appliquons le chapeau haut de forme de l'image inversée, une seule structure est détectée qui correspond à l'ensemble des pixels du vélo (voir Figure 11.2(b)). Il est donc nécessaire d'inverser uniquement les pixels différents de zéro. Le nouveau résultat de détection, présenté par la Figure 11.2(c), illustre la bonne détection du vélo. Par contre, la voiture reste liée au bord et elle ne sera pas détectée. En ajoutant le bord du côté de la rue, la voiture est finalement détectée, comme le montre la Figure 11.2(d).

En fin, un seuil de  $10cm$  est appliqué à l'image résultante pour éliminer les artefacts produits par la rugosité du sol et les surfaces bruitées. La méthode est résumée par le diagramme de la Figure 11.3(a).

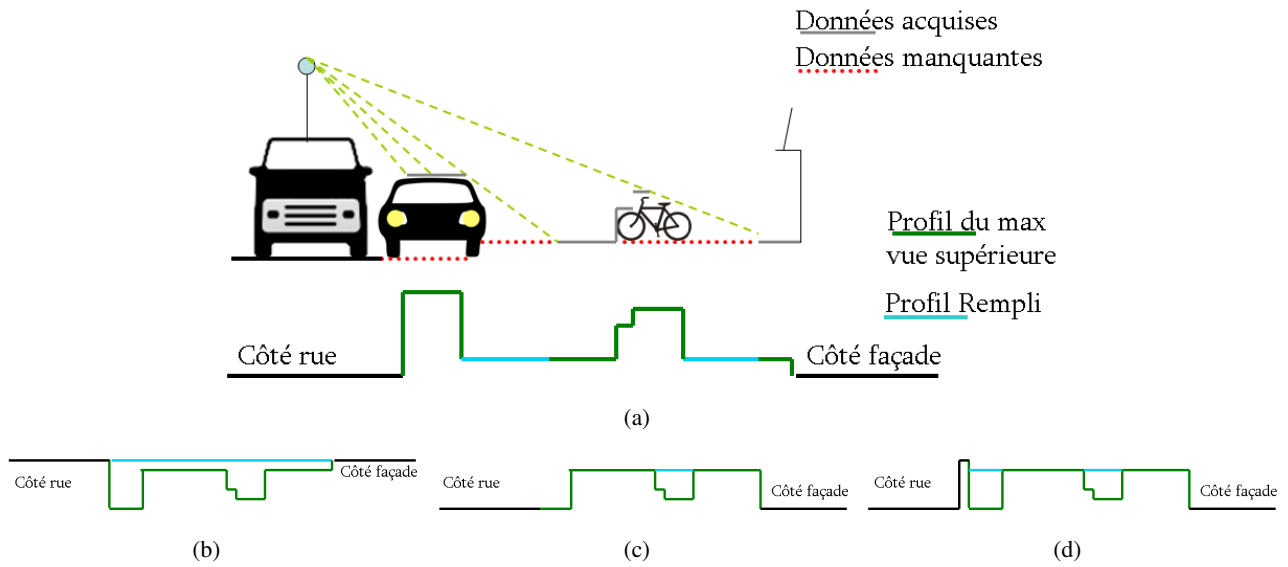


FIGURE 11.2 – (a) Illustration d'un artefact proche du capteur. (b) Problème d'inversion de points à zero. (c) Problème des artefacts collés au bord . (d) Résultat final.

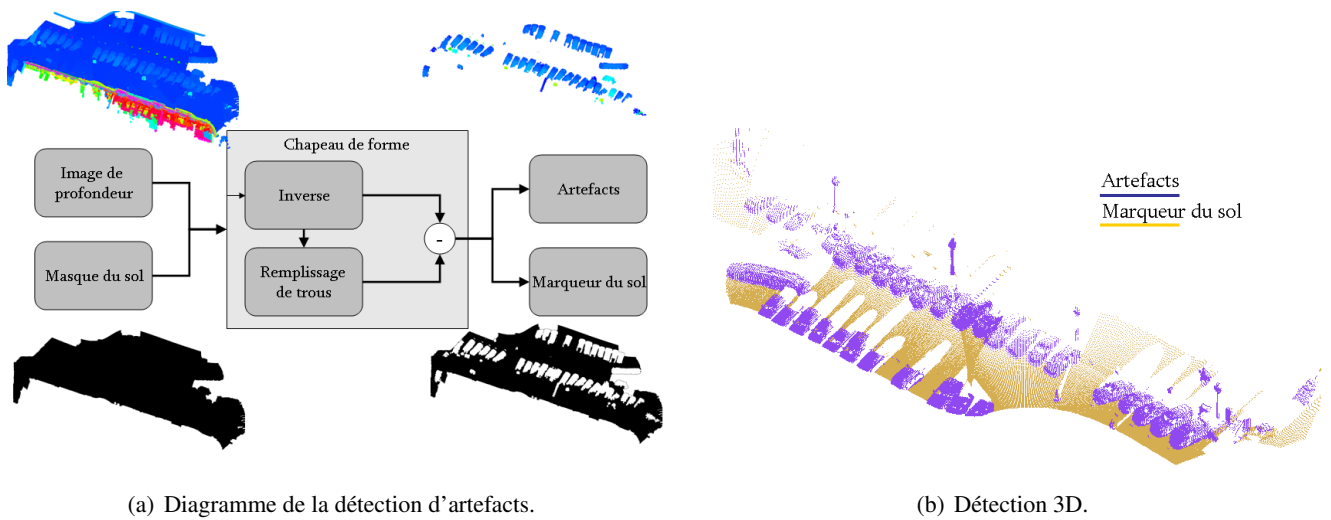


FIGURE 11.3 – Les étapes de la détection d'artefacts.

La méthode présente deux sorties : l'image d'artefacts et une image que nous appellerons le marqueur du sol. Comme on peut l'observer, ce marqueur est le même que le masque initial sans les pixels qui appartiennent aux artefacts. Les images de sortie peuvent être rétro-projetées sur les nuages de points pour réaliser la segmentation 3D comme la Figure 11.3(b) l'illustre.

## 11.3 Classification d'artefacts

### 11.3.1 Artefacts connectés

Les artefacts sont bien détectés, ce qui permet de filtrer correctement les données pour une modélisation des façades et du sol. Par contre, leur classification nécessite que chaque artefact soit une composante connexe ( $CC$ ) séparée des autres. Cependant, en utilisant une simple labellisation, il arrive que plusieurs artefacts restent connectés entre eux, formant une seule  $CC$ . La Figure 11.4 illustre plusieurs situations d'une  $CC$  avec plusieurs artefacts : des piétons en train de toucher un lampadaire (Figure 11.4(a)), deux voitures trop proches (Figure 11.4(b)) et un piéton en train d'ouvrir la porte de sa voiture (Figure 11.4(c)).

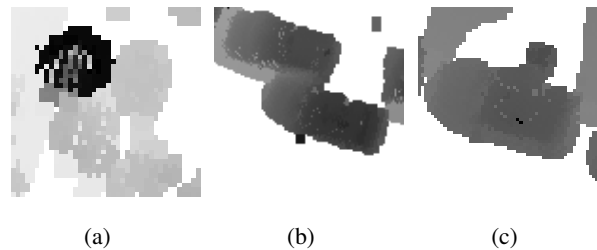


FIGURE 11.4 – Exemples d'artefacts connectés.

Nous avons décrit un artefact comme une bosse, ainsi le nombre d'artefacts connectés dans une même  $CC$  dépend du nombre de bosses qu'elle renferme. Ces bosses peuvent être exprimés en termes de nombre de maxima de la  $CC$ . Cependant, une  $CC$  peut avoir plusieurs maxima, causés par de petits pics et de la rugosité dans la partie supérieure de l'artefact. Pour extraire les maxima les plus significatifs, deux filtres morphologiques sont utilisés. Une ouverture par surface [164] et un opérateur  $h$ -Maxima [150]. L'opérateur  $h$ -Maxima élimine les maxima qui ont une profondeur inférieure ou égale à un seuil  $h$ . Il utilise une reconstruction morphologique par dilatation  $R_f^\delta(g)$  où le marqueur ( $g$ ) est une image à laquelle on a soustrait une constante  $h$  de  $10cm$ .

Une fois les artefacts filtrés, on utilise la ligne de partage des eaux de l'image de gradient avec des marqueurs. Les marqueurs sont les maxima et le fond de l'image filtrée. La Figure 11.5 illustre trois exemples de la séparation d'artefacts connectés, en montrant la performance de la méthode.

### 11.3.2 Analyse de caractéristiques

Chaque  $CC$  de l'image est analysée afin de déterminer le type d'artefact. Plusieurs catégories d'artefacts sont considérées : des lampadaires, des voitures et des piétons. Une quatrième classe qui reprend le reste (des poubelles, des motos, des parcs mètres, des abris de bus, etc.) est appelée "Autres".

Les mesures calculées sont :



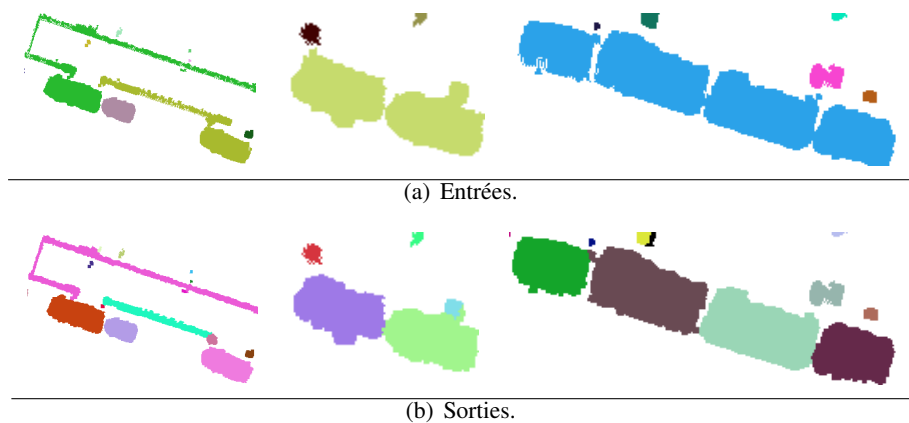


FIGURE 11.5 – (a) Détection initiale et (b) Séparation d'artefacts.

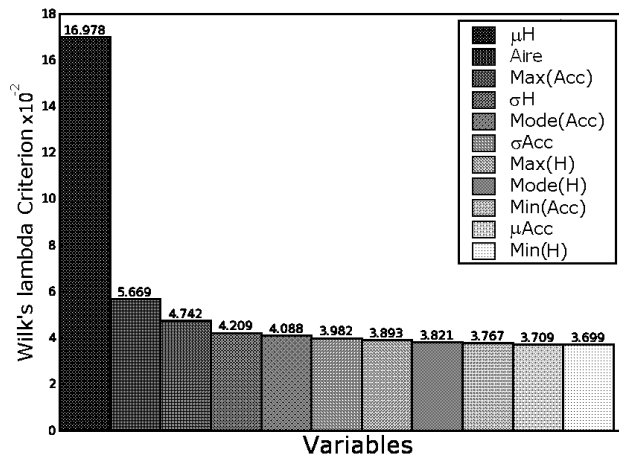
1. La moyenne, l'écart type, le maximum, le minimum et le mode de la hauteur [mètres] ;
2. La moyenne, l'écart type, le maximum, le minimum et le mode de l'accumulation [nombre de points] ;
3. La surface de la composante connexe. Initialement, cette caractéristique est calculée en pixels. Toutefois, en utilisant l'information de la caméra, la valeur est transformée en mètres carrés. Cette transformation homogénéise les éléments qui appartiennent aux mêmes classes dans l'ensemble d'images (les images de profondeur sont générées de manière indépendante, selon la résolution du système d'acquisition).

Nous avons manuellement annoté une base de données pour établir une vérité terrain de 442 artefacts : (67 voitures, 33 lampadaires, 198 piétons et 144 autres). Afin d'analyser les caractéristiques par rapport à leur pouvoir discriminant nous utilisons le critère de lambda de Wilk [101, 135]. Le Lambda de Wilk est une statistique utilisée par l'analyse factorielle discriminante [10] dans le but de tester si plusieurs groupes d'observations multi-variées ont des moyennes significativement différentes. L'idée est de sélectionner les variables au fur et à mesure en utilisant le critère de lambda. La Figure 11.6(a) montre les variables ordonnées selon cette méthode.

Une distribution de classes avec les trois premières caractéristiques est illustrée sur la Figure 11.6(b). Ici, la surface est la variable caractéristique des voitures et la moyenne de la hauteur celle des lampadaires. Néanmoins, la classe "Autres» est entrecroisée principalement avec la classe des piétons. Ce comportement est mis en évidence dans le processus de classification de la section 11.3.3.

### 11.3.3 Classification

En utilisant les onze caractéristiques, nous éliminons d'abord les *CCs* de petite taille ( $1cm^2$ ) et dont le maximum de l'accumulation est inférieur à 3 points. Puis nous utilisons deux méthodes de classification supervisées, l'analyse linéaire discriminante (ALD) [43] et les séparateurs à vastes marges (*support vector machine SVM*) [27, 35, 160]. La classification a été validée en utilisant la validation croisée (*K-fold cross validation*,  $K=10$ ).



(a) Sélection de caractéristiques (Critère ordonné de lambda de Wilk

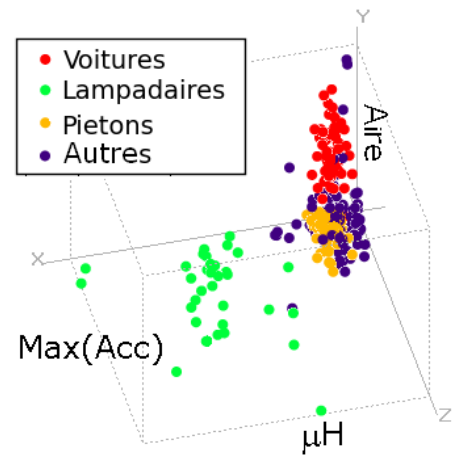

 (b) Distribution de classes  $\mu H$  vs Surface vs max Acc

FIGURE 11.6 – Sélection de caractéristiques.

Les résultats de classification sont présentés dans les matrices de confusion des Tables 11.1 et 11.2, respectivement. Les tables montrent de bons résultats de classification avec l'utilisation des onze caractéristiques. Le classifieur SVM présente une meilleure performance que l'ALD, car l'erreur totale de classification en utilisant l'ALD est de 18,09% et de 14,93% en utilisant le SVM. Dans le cas du SVM, tous les lampadaires ont été bien classifiés, et les voitures et les piétons ont de bons taux de classification, de 91,04% et de 96,97%. Néanmoins, une grande quantité d'artefacts de la classe "Autres» ont été classifiés comme des piétons. Un exemple de ce problème est présenté avec les horodateurs qui sont des structures avec des caractéristiques équivalentes à celles des piétons (hauteur et surface similaires). Les deux classifieurs présentent ce problème d'entrecroisement.

Vérité Terrain	Détectés			
	Voitures	Lampadaires	Piétons	Autres
<b>Voitures</b>	87,08%	0,0%	1,35%	11,55%
<b>Lampadaires</b>	0,0%	90,42%	9,58%	0,0%
<b>Piétons</b>	0,0%	0,0%	97,16%	2,83%
<b>Autres</b>	6,82%	0,06%	39,90%	53,20%

TABLE 11.1 – Matrice de Confusion ALD.

Un exemple de détection et classification d'artefacts est présenté dans la Figure 11.7. Comme la Figure 11.7(a) l'illustre, la détection d'artefacts (même sans classification) est utile pour filtrer les données et faciliter la modélisation

	Détectés			
Vérité Terrain	Voitures	Lampadaires	Piétons	Autres
Voitures	91,04%	0,0%	0,0%	8,96%
Lampadaires	0,0%	100,0%	0,0%	0,0%
Piétons	0,0%	0,0%	96,97%	3,03%
Autres	7,64%	1,39%	28,47%	62,50%

TABLE 11.2 – Matrice de Confusion SVM.

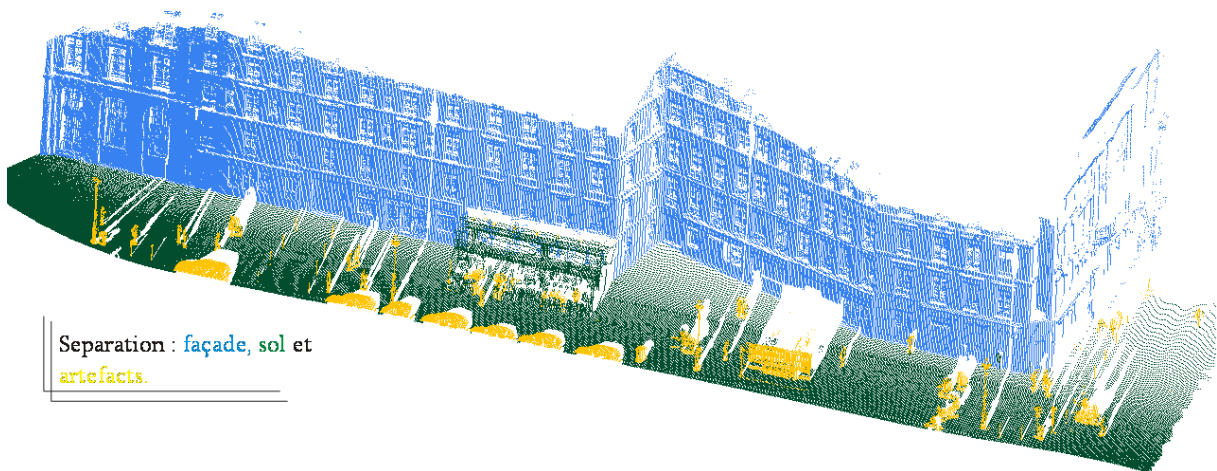
des façades et du sol. La méthode permet la détection de lampadaires, de voitures et de piétons ainsi que la localisation d'autres artefacts non encore reconnus (Figure 11.7(b)). Cependant dans cet exemple, un abri de bus a été détecté dans la classe des lampadaires et une petite voiture dans la classe "Autres».

## 11.4 Discussion

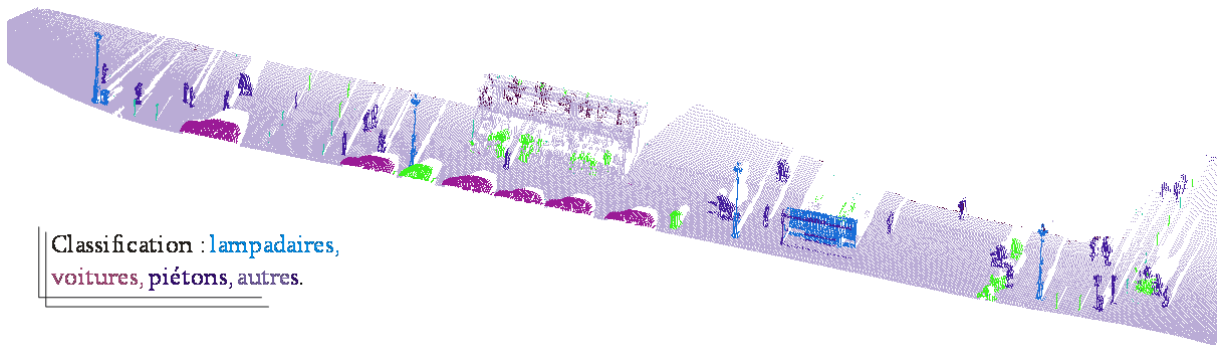
Dans ce chapitre nous avons ouvert une voie d'exploration pour l'extraction et la classification d'artefacts. La méthode de détection est basée sur le chapeau haut de forme du remplissage de trous. La méthode détecte, de manière automatique et avec les mêmes paramètres, les artefacts d'une scène quelle que soit leur forme. Une fois que les artefacts sont détectés, et sans avoir besoin de la classification, le filtrage des données de façades et du sol devient possible. Cette méthode présente de bons résultats pour l'ensemble de la base de données. Cependant, quelques artefacts échappent à la détection car ils sont trop fins ou proches à la façade. Cette dernière situation se produit du fait que les artefacts sont connectés au bord du masque du sol (côte façade) et par conséquent le chapeau haut de forme ne les détecte pas.

Le filtrage de données du sol est fait en utilisant la rétro-projection de l'image d'artefacts et le masque du sol sur le nuage de points. Néanmoins, des points du sol acquis par le capteur sont assignés à des points d'artefacts parce que la détection est faite à partir d'une vue supérieure de la scène et que le sol se trouve en-dessous de l'artefact. Si un artefact est non-convexe (comme les arbres) le nombre de points du sol acquis assignés à l'artefact peut être élevé. La Figure 11.8 schématise la détection de deux artefacts : une voiture et un arbre. La voiture est plutôt convexe et c'est pour cela que le télémètre laser n'acquiert pas de données de sol sous la voiture. Par contre, l'arbre est non-convexe et plusieurs points du sol sont acquis. Après la détection, ces points sol qui se trouvent en-dessous des artefacts seront assignés par erreur comme points de l'artefact. Un traitement en 3D peut être mis en place pour la récupération de ces points comme données de sol, qui sont nécessaires dans la segmentation de la chaussée (voir Chapitre 12).

Dans la section 11.3.1 nous avons présenté une méthode pour la séparation d'artefacts qui sont connectés dans une même composante connexe  $CC$ . Cette séparation est un point important et nécessaire en vue d'aboutir à une bonne



(a) Détection d'artefact et Filtrage de données façade-sol.



(b) Classification d'artefacts.

FIGURE 11.7 – Résultats des méthodes de détection et classification d'artefacts.

classification des artefacts. Néanmoins, à partir d'une vue supérieure, plusieurs artefacts peuvent être masqués par d'autres et ne pourront donc être séparés. Par ailleurs, la méthode est basée sur le nombre de maxima dans la  $CC$ , étant donné qu'un artefact est décrit comme une bosse depuis la vue supérieure. Cependant, certains artefacts (les arbres par exemple) ont plusieurs maxima et la  $CC$  sera découpée en plusieurs artefacts. Ces deux problèmes seront illustrés plus en détail dans les sections 11.4.1 et 11.4.2, respectivement.

Par rapport à la classification, deux méthodes d'apprentissage supervisé ont été testées, ALD et SVM. Le meilleur taux de classification est obtenu avec le classifieur SVM avec un taux d'erreur de 14,93%. Le principal problème dans la classification est la grande variabilité intra-classe de la classe "Autres" qui implique trop d'erreurs. Une analyse approfondie de cette classe est nécessaire afin d'augmenter la performance de la classification. Par ailleurs, la classification est basée sur l'extraction de caractéristiques de l'image de profondeur de la  $CC$  de chaque artefact. C'est pourquoi, avec cette méthode, on ne pourra pas distinguer un piéton d'un horodateur car ils ont des

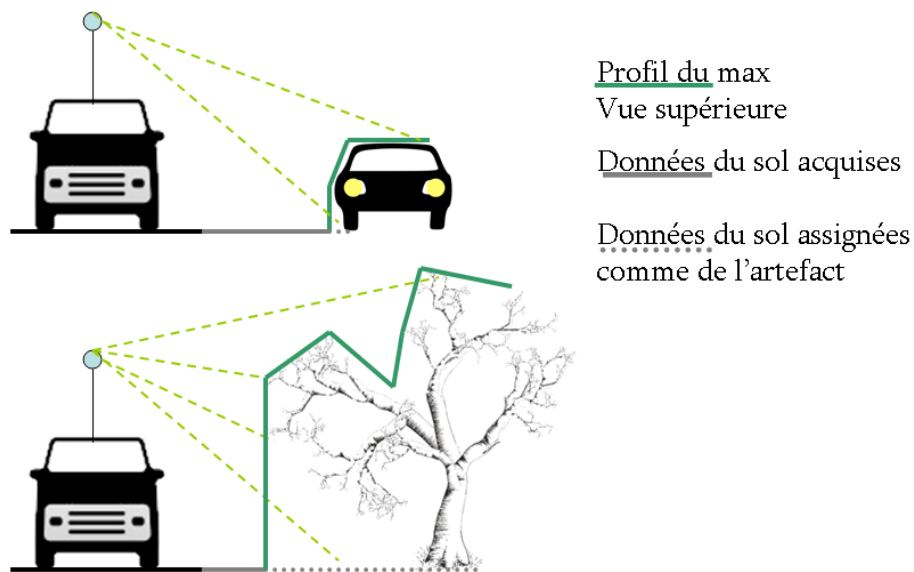


FIGURE 11.8 – Points sol assignés comme artefact.

caractéristiques très similaires. Plusieurs alternatives sont envisageables pour améliorer la détection et la classification dans ce type de circonstances :

- Pour une première détection d'artefacts fixes et mobiles, l'acquisition en deux temps de la scène est une option, sous la contrainte d'avoir besoin du double de mémoire de stockage et de temps de calcul.
- L'utilisation de l'information couleur et des descripteurs de forme 3D comme cela a été présenté dans [121] pour la détection de voitures. Afin d'utiliser un processus multimodal, couleur et géométrie, les données doivent être soigneusement recalées, et de plus acquises dans le même instant de temps. Du côté des descripteurs de forme 3D [83], une vérité terrain est nécessaire pour chaque artefact à détecter. Ce point rend la tâche plus difficile dans le cas de plusieurs artefacts.
- Dans le cas de données 3D issues d'un processus de stéréogrammétrie, l'information 2D et 3D est facilement exploitable pour la détection d'artefacts comme cela a été, par exemple, présenté dans [22]. Cette approche présente une détection d'artefacts (des voitures, des piétons, des arbres, des panneaux de signalisation, des barrières) et d'autres éléments de la scène urbaine (des bâtiments, des trottoirs, des rues) en utilisant une vérité terrain de 600 images segmentées manuellement de frames de vidéo. La méthode exploite les caractéristiques couleurs de l'image, le mouvement de la vidéo et le 3D de la stéréo. La méthode des forêts aléatoires comme technique de classification est utilisée.

### 11.4.1 Artefacts cachés et non-séparation d'artefacts

Comme nous l'avons déjà mentionné, la détection d'artefacts est réalisée en utilisant une projection sur images de la vue supérieure de la scène. Pourtant, certains artefacts de grande taille peuvent occulter d'autres artefacts qui se trouvent en-dessous.

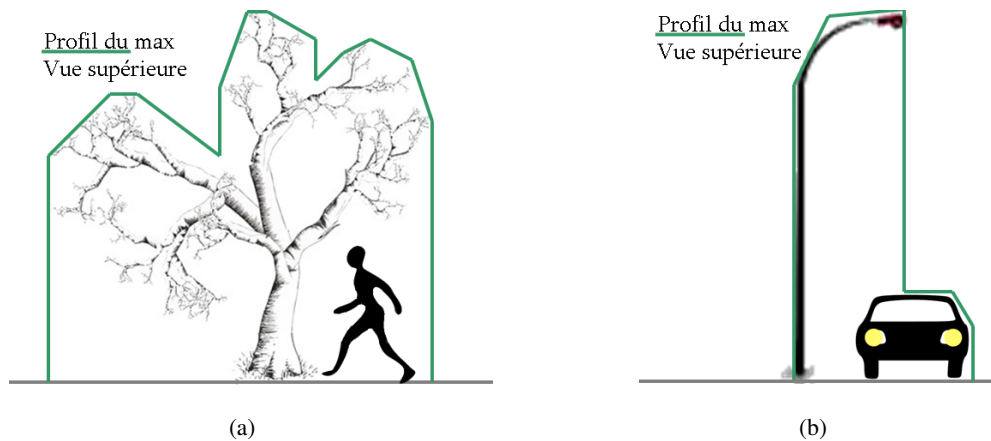


FIGURE 11.9 – Illustration d'artefacts cachés.

Un exemple de cette problématique est illustrée sur la Figure 11.9(a). Avec notre méthode, le piéton ne sera pas détecté et ses points 3D seront assignés à l'arbre. La Figure 11.9(b) montre d'ailleurs un exemple où même si l'occultation n'est pas totale, les deux artefacts restent connectés dans la même  $CC$  car celle-ci n'a qu'un maximum. Des méthodes d'analyse 3D, par exemple la reconnaissance d'objets dans des points 3D [83], pourront être utilisées comme étape de post-traitement pour la séparation d'artefacts.

### 11.4.2 Non-séparation d'artefacts

On vient de présenter la non-séparation comme un problème qui réduit la performance de notre méthode. Néanmoins, la séparation indésirable de points 3D d'un même artefact produit une caractérisation erronée et donc une mauvaise classification.

La Figure 11.10 montre ce problème, car l'arbre est séparé en trois parties. Une solution envisageable est de séparer une  $CC$  seulement si le nombre de maxima est inférieur à une valeur donnée. Dans le cas concret des arbres, nous pourrions également mesurer la variabilité ou l'entropie de la forme et de la hauteur de la  $CC$ . Cette information est discriminante pour les arbres du fait de leur haute variabilité.

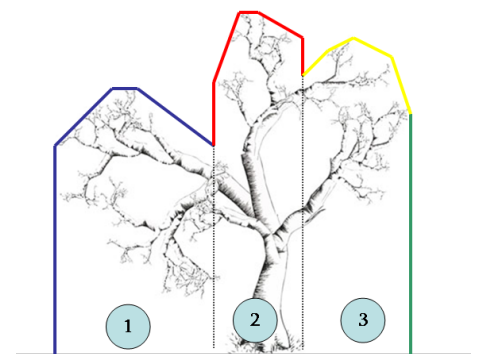


FIGURE 11.10 – Illustration d'une séparation indésirable.

## Segmentation de la Chaussée

### Sommaire

<b>12.1 Introduction</b>	<b>163</b>
<b>12.2 Projection sur images sans artefacts</b>	<b>164</b>
<b>12.3 Segmentation des points du sol</b>	<b>165</b>
12.3.1 Sélection de marqueurs	166
12.3.2 Estimation du gradient	167
<b>12.4 Sélection de frontières des régions</b>	<b>169</b>
<b>12.5 Discussion</b>	<b>170</b>

### 12.1 Introduction

La plupart des approches dans la modélisation d'environnements urbains se focalisent sur la segmentation et la modélisation de bâtiments et de façades. Néanmoins dans l'objectif d'une modélisation complète de l'environnement urbain, l'information de la chaussée est aussi importante. On trouve dans la Figure 12.1 trois exemples de diverses étapes de la modélisation urbaine dont la première consiste à modéliser un bâtiment (voir Figure 12.1(a)). En utilisant la méthode de détection et de classification des artefacts, présentée dans le chapitre précédent, les lampadaires peuvent être ajoutés à la modélisation de la scène, comme l'illustre la Figure 12.1(b). Cependant, la scène reste incomplète, l'information de la chaussée est nécessaire afin de perfectionner la modélisation et d'avoir une vue d'ensemble. Le modèle complet avec la modélisation du trottoir et de la rue est présenté dans la Figure 12.1(c) dont nous pouvons remarquer le réalisme de la scène en la comparant au modèle minimaliste d'un simple bâtiment de la Figure 12.1(a).

Dans les données issues des systèmes d'acquisition, principalement les données LIDAR, on trouve l'information de la chaussée. Ces données nous permettent d'exploiter celle-ci et donc d'alimenter le modèle urbain en augmentant son réalisme.

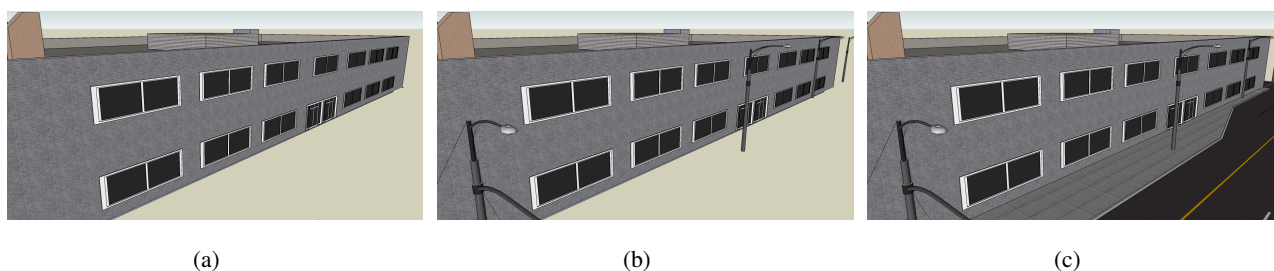


FIGURE 12.1 – Modélisation : (a) d'un bâtiment, (b) avec des lampadaires et (c) avec la chaussée.

De manière générale, les approches de segmentation de routes ont pour but la détection des axes routiers et la délimitation de zones urbaines en utilisant des données aériennes LIDAR [166] et des données images [71]. Un état de l'art de ces types d'approches a été présenté dans [136]. Par ailleurs, avec des données prises au sol, où la séparation trottoir-rue devient possible, les travaux de recherche ne sont pas nombreux. Brostow *et al.* [22] cherchent la segmentation des pixels vidéo et des points 3D, issus de la stéréogrammétrie ; répartis en douze classes parmi lesquelles on rencontre le trottoir et la rue. De plus, une large base de données d'images étiquetées manuellement a été en 2009 créée [82]. La base possède les étiquettes route et trottoir. Cependant, les travaux qui exploitent cette base se focalisent sur la détection de bâtiments et de fenêtres [81].

Dans ce chapitre nous présentons une méthode de segmentation de la chaussée à partir de données LIDAR de systèmes mobiles. La méthode s'appuie sur la projection du minimum sur images de données sol en réduisant les problèmes des artefacts (voir section 12.2). Cette image est segmentée en utilisant l'algorithme de zones quasi plates comme l'illustre la section 12.3. La section 12.4 présente la détection de la séparation trottoir-rue avec l'aide d'une représentation en un graphe.

## 12.2 Projection sur images sans artefacts

Les principaux problèmes dans la segmentation automatique de la chaussée sont les artefacts et les occlusions produites durant l'acquisition des données. La Figure 12.2(a) montre un exemple où tous les points du sol sous les arbres ont été manqués.

De plus, en utilisant la projection de la distance maximale à partir de la vue supérieure, plusieurs points du sol sont masqués par les artefacts (voir Figure 12.3(a)). Nous pourrions penser qu'en projetant les points avec la distance minimale, ce problème serait résolu. Cependant, ce n'est pas le cas parce que plusieurs points des artefacts n'ont pas leurs points au sol correspondants. Ainsi, en projetant la distance minimale, ces points artefacts y resteront comme l'illustre la Figure 12.3(b). En comparant les Figures 12.3(a) et 12.3(b) nous observons une diminution considérable du nombre de pixels d'artefacts dans l'image, néanmoins celle-ci n'est pas suffisante pour une analyse



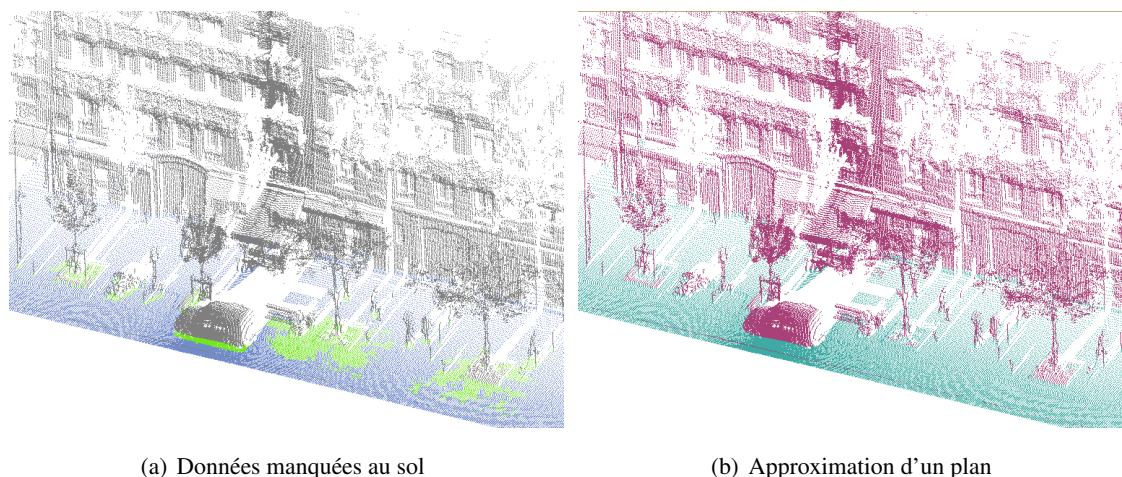


FIGURE 12.2 – (a) Points du sol manqués en vert et (b) Points qui appartiennent au plan estimé en cyan.

au sol car les artefacts produisent encore des discontinuités à ce niveau. Afin de récupérer ces points du sol sans les artefacts, nous rétro-projetons le marqueur du sol issu de l'étape de détection d'artefacts (voir Figure 11.3(a)). Ensuite, une approximation d'un plan avec les points sélectionnés est réalisée. Les points qui se trouvent à une distance inférieure à une valeur donnée seront considérés comme points du sol. La distance choisie est deux fois l'écart-type de la distance entre le plan estimé et les points utilisés dans l'approximation, et de cette manière, nous obtenons les points du sol acquis par le télémètre laser. La Figure 12.2(b) illustre la sélection des points du sol. En utilisant les points sélectionnés, nous régénérons une nouvelle projection sur images avec la distance minimale des points au plan de la caméra. Nous utilisons la distance minimale pour prendre principalement l'information qui se trouve dans la partie basse de la scène.

Finalement, nous filtrons l'image avec une ouverture par surface de taille  $50\text{cm}^2$ , afin d'éliminer certains morceaux d'artefacts qui sont restés dans la projection. Une comparaison des images utilisées pour la détection d'artefacts et pour la segmentation de la chaussée est illustrée sur la Figure 12.3. On peut observer que l'information au-dessous des arbres a été récupérée. D'ailleurs nous voyons, en comparant les Figures 12.3(c) et 12.3(d), que plusieurs points de bruit ont été enlevés car les structures du sol deviennent visibles à l'œil nu.

### 12.3 Segmentation des points du sol

Le but de cette segmentation est de séparer les pixels du sol qui appartiennent au trottoir et à la rue en régions disjointes. La méthode utilise la ligne de partage des eaux contrainte par des marqueurs. La détection des marqueurs, nécessaire pour la segmentation, est présentée dans la section 12.3.1. La Figure 12.4 illustre le processus en général et un exemple de segmentation.

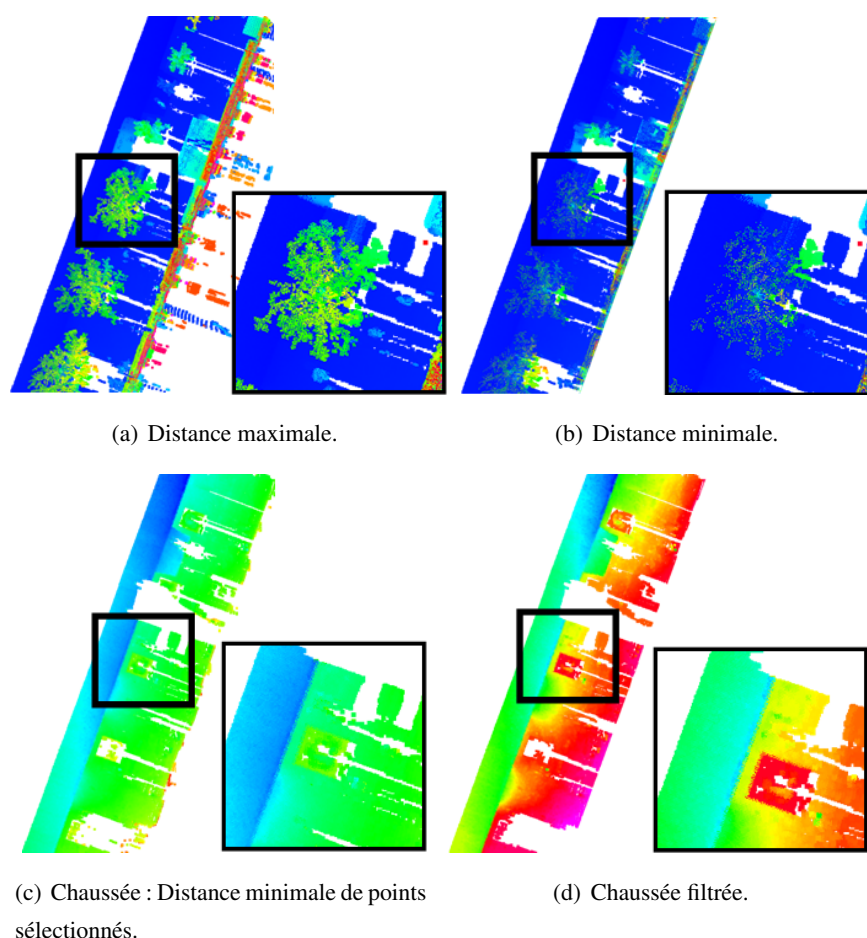


FIGURE 12.3 – Projection sur images.

### 12.3.1 Sélection de marqueurs

D'abord, une segmentation fine de l'image de profondeur est réalisée en utilisant l'algorithme de  $\lambda$ -zones plates (voir Définition 19). L'idée de cette segmentation est d'avoir le nombre minimum de régions, tout en gardant les régions de la route et du trottoir séparées, et c'est pourquoi le choix de la valeur  $\lambda$  est important. La valeur de  $\lambda$  doit être inférieure à la hauteur minimale qui sépare la rue du trottoir. Malheureusement, dans l'aménagement urbain, les normes strictes de la hauteur de la chaussée ne sont pas disponibles. Néanmoins, dans les zones métropolitaines, les hauteurs proposées sont :  $20mm$  pour les passages piétons (accès handicapés),  $40mm$  pour l'accès des véhicules (parkings principalement) et  $140mm$  pour le trottoir.

La Figure 12.5 montre un exemple de segmentation avec deux valeurs de  $\lambda$  inférieures à  $20mm$ . Cependant, comme nous pouvons l'observer dans la Figure 12.5(d), même avec une valeur de  $10mm$ , plusieurs régions du trottoir sont fusionnées avec la rue. C'est pour cela qu'après plusieurs tests, une valeur de  $5mm$  a été choisie afin d'éviter

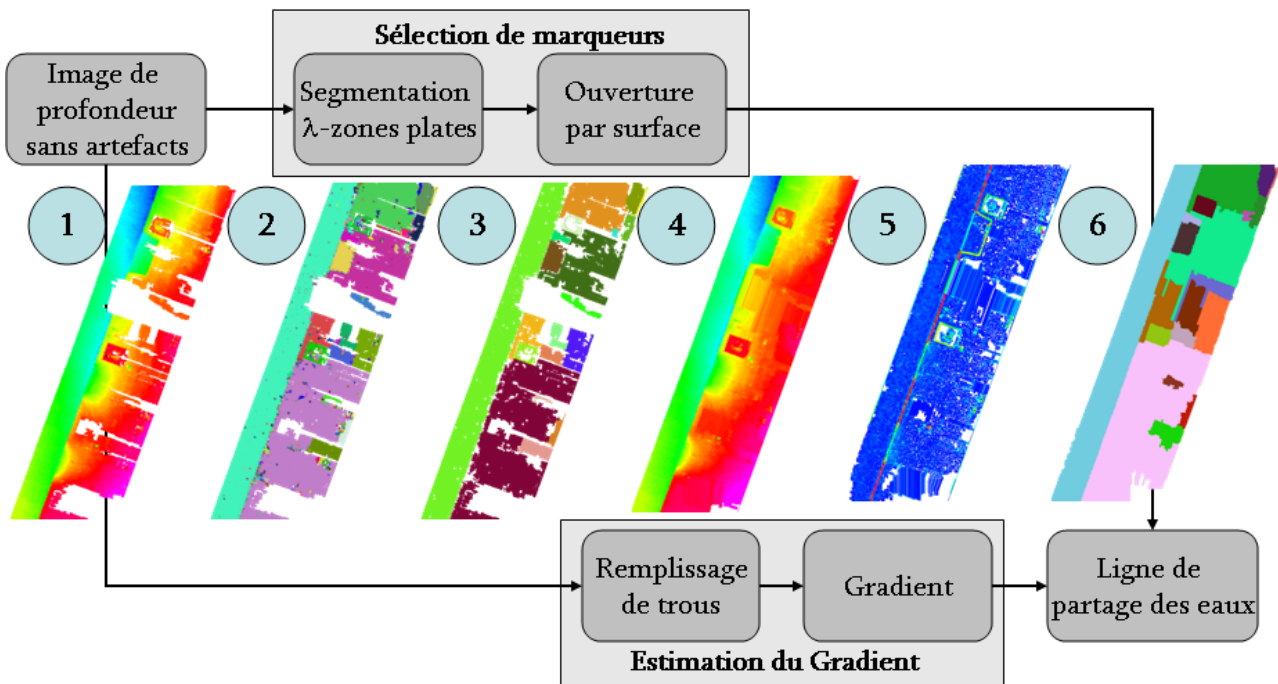


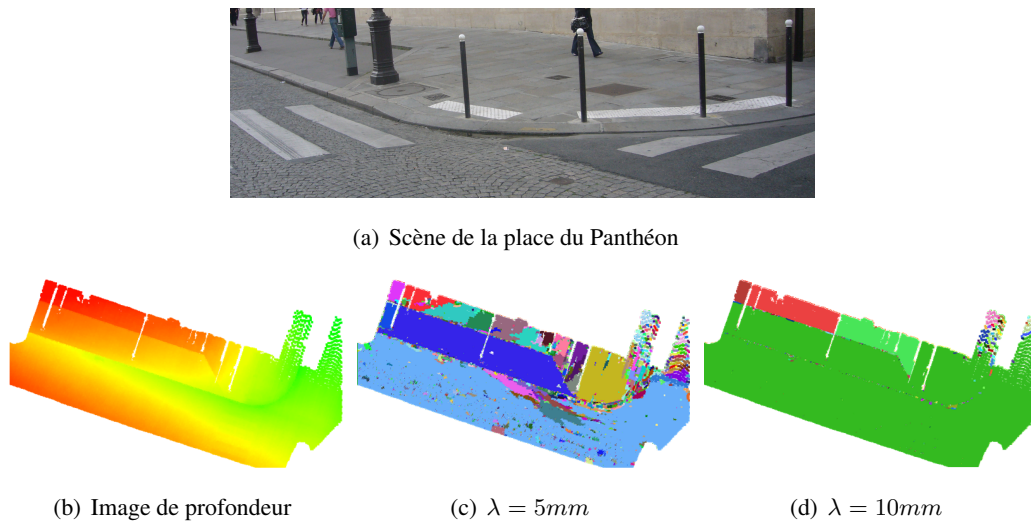
FIGURE 12.4 – Diagramme de la segmentation de la chaussée.

les fuites vers les endroits d'accès aux personnes à mobilité réduite. Le résultat de la segmentation  $5mm$ – zones plates est illustré sur la Figure 12.5(c).

Avec cette valeur de  $\lambda = 5mm$  la méthode produit beaucoup de petites régions, à cause de la rugosité du sol. En conséquence, une sélection par attribut de surface est effectuée, en choisissant les régions de plus de  $20cm^2$  comme marqueurs de la segmentation.

### 12.3.2 Estimation du gradient

Le calcul du gradient est fait en utilisant un gradient morphologique de l'image de profondeur. Néanmoins, une étape préalable pour estimer l'information manquante produite par les occlusions d'artefacts est mise en œuvre. Dans la section 10.4, nous avons présenté une méthode pour le remplissage de trous afin de séparer la façade du sol. Dans cette méthode une propagation des minima est réalisée de manière locale sans tenir compte des différentes variations du sol. Cette approximation est suffisante pour estimer grossièrement l'information du sol et donc le séparer de la façade. Par contre, dans le cas de la séparation trottoir-rue, cette estimation pourrait donner des résultats erronés. Une illustration d'un trou dans la séparation trottoir-rue est présentée sur la Figure 12.6(a). En appliquant le remplissage de la Définition 20 nous obtenons la Figure 12.6(b). Idéalement, le remplissage d'un trou doit suivre d'un côté l'information de la rue et d'un autre côté l'information du trottoir comme l'illustre la Figure

FIGURE 12.5 – Segmentation de  $\lambda$ -zones plates.

12.6(c).

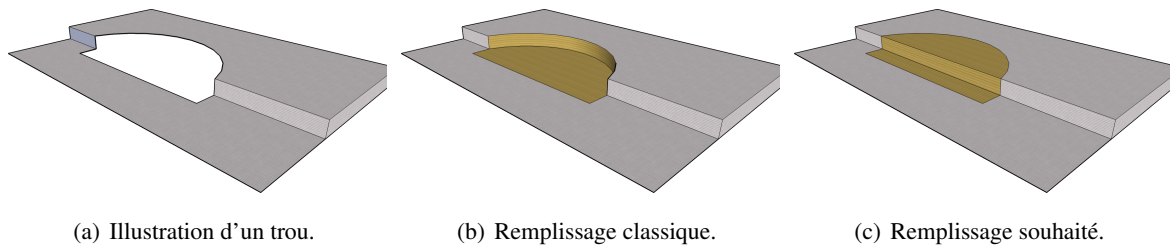


FIGURE 12.6 – Illustration de l'information manquante dans la séparation trottoir-rue.

Après plusieurs observations de la zone de test, nous avons constaté que la direction de la ligne de séparation trottoir-rue coïncide avec la direction de la façade. La détection de la ligne de la façade est réalisée en utilisant la transformée de Hough qui a été décrite par la section 10.3.1. Pour l'exemple de la Figure 12.2, la ligne de la façade est illustrée dans la Figure 12.7(a). Une fois la ligne détectée, l'information manquante est estimée dans la famille de lignes parallèles à la ligne de la façade. Le résultat du remplissage de la Figure 12.7(b) est présenté dans la Figure 12.7(c). Nous pouvons constater la bonne estimation de la délimitation trottoir-rue et des autres structures comme la région parking et la délimitation de la partie basse des arbres. Dans le zoom 2 de la Figure 12.7(c), nous remarquons certaines traces dans la direction de la façade car la rue est en pente et les trous ont une dimension importante. C'est pour cette raison que nous estimons l'information manquante pour calculer le gradient et non pas pour extraire les marqueurs de la segmentation.

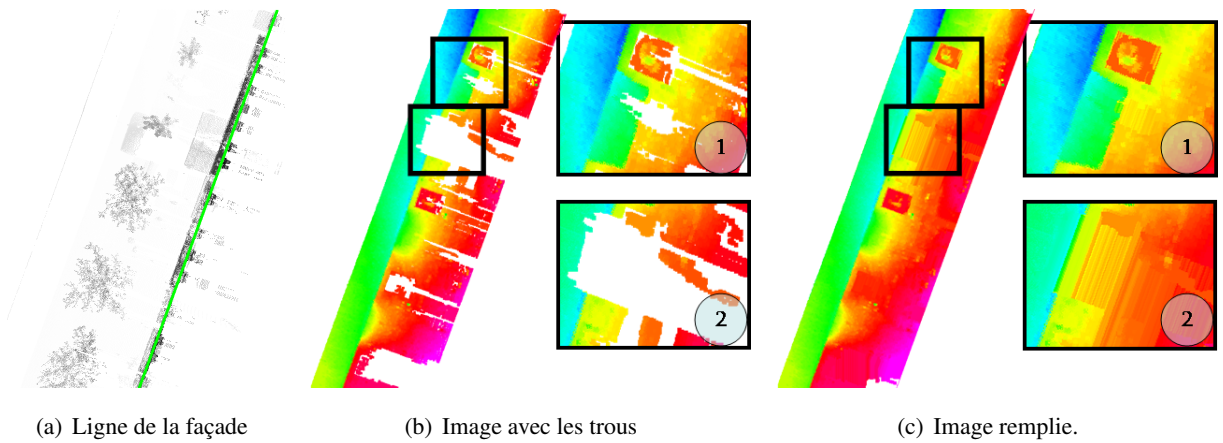


FIGURE 12.7 – Remplissage directionnel des trous en utilisant la ligne de la façade. (a) La ligne de la façade en vert.

## 12.4 Sélection de frontières des régions

Une fois que la segmentation est calculée, un graphe d'adjacence des régions (*region adjacency graph RAG* en anglais) est obtenu pour représenter ce résultat. La représentation *RAG* permet de sélectionner les régions frontalières entre le trottoir et la route.

Notons *RAG* comme  $G = (\mathcal{R}, E, W)$  où  $\mathcal{R} = \{r_i\}$  est une partition en régions disjointes et  $E = \{e_{ij} = r_i, r_j\}$  est l'ensemble des arêtes représentant les relations de voisinage. A chaque arête  $e_{ij}$ , un poids  $w_{ij}$  peut être assigné. Ce poids donne une mesure de la différence entre les deux régions. Dans notre cas, le poids  $w_{ij}$  est simplement obtenu à partir de la différence moyenne calculée de part et d'autre de la frontière sur un bord d'épaisseur  $d$  (voir la Figure 12.8). Ce calcul du contraste d'un bord augmente la robustesse de la méthode contre les accès des piétons et la rugosité du sol. Après plusieurs tests, nous avons choisi  $d = 5pix$ . Finalement, les arêtes d'une valeur entre  $50mm$  et  $180mm$  sont choisies comme la frontière entre le trottoir et la route.

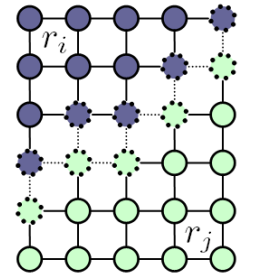


FIGURE 12.8 – Le bord local de l'épaisseur  $d = 1$ .

La Figure 12.9 illustre le bord local pour calculer  $w_{ij}$ , la génération du *RAG* et la segmentation finale (extraction du contour). On peut voir la bonne délimitation de la chaussée. De plus, on observe la détection de la zone de parking et le bord de la plantation des arbres.

Dans la même dynamique que tous nos traitements de nuages de points sur images, présentés dans les chapitres précédents, les résultats de la segmentation de la chaussée sont rétro-projetés sur les nuages de points. Un exemple de cette segmentation 3D est illustré sur la Figure 12.10.



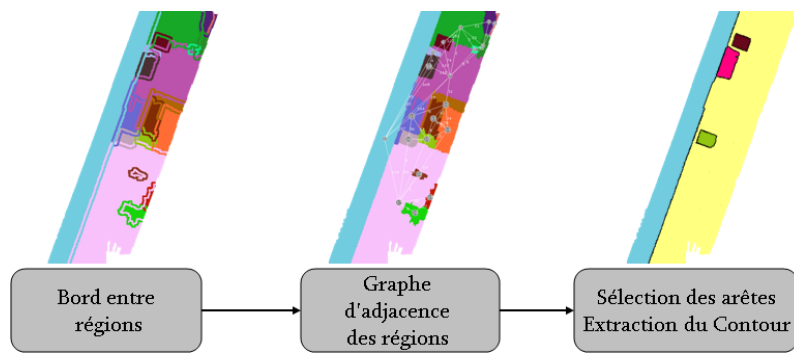


FIGURE 12.9 – Diagramme de la segmentation finale de la chaussée.

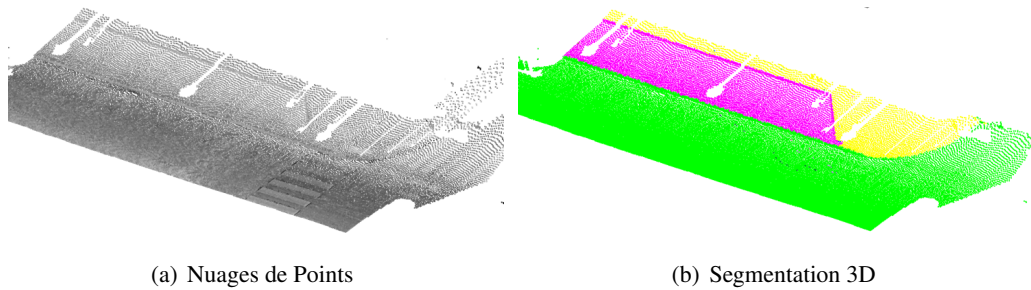


FIGURE 12.10 – Projection de la segmentation sur le nuage de points original

## 12.5 Discussion

Dans ce chapitre nous avons présenté une méthode de détection de la séparation de la rue et du trottoir. La méthode utilise la projection sur des images de la valeur minimale de données du sol pour réduire l'influence des artefacts. Les données au sol sont sélectionnées en utilisant l'approximation au plan du marqueur du sol. Cette approximation est une stratégie simple de mise en œuvre afin de récupérer les points 3D cachés sous des artefacts (Par exemple un arbre) qui n'ont pas été détectés par la segmentation à partir de la vue supérieure. Cependant, un plan n'est pas l'estimation optimale quand il s'agit des formes de chaussée plus complexes, telles que les intersections de rues. C'est pourquoi nous proposons d'explorer une approximation avec des courbes paramétriques, par exemple les splines ou les NURBS [124]. Ces courbes permettront à la fois de sélectionner les points du sol et d'estimer les données manquantes produites par les occlusions.

Un autre point important à mentionner est la précision de notre méthode. D'abord la méthode présentée est seulement utilisée avec les données issues de systèmes mobiles, car ceux-ci ont une haute résolution de points au niveau de la chaussée. En effet, nous travaillons dans les limites de la résolution et de la précision du système car l'al-

gorithme de zones quasi plates utilise une valeur de  $\lambda$  de  $5mm$ . C'est pour cela que les problèmes d'acquisition, comme les sauts du GPS, affecteront de manière considérable la segmentation. La méthode est robuste par rapport aux petits points de bruit car ils sont éliminés après la sélection des régions les plus grandes. En outre, nous avons bien segmenté les accès des piétons ( $20mm$ ) grâce à leur représentation sur un graphe. Toutefois, si la hauteur de l'accès des piétons est plus basse que  $20mm$ , une partie du trottoir peut être fusionnée avec la route, produisant une mauvaise segmentation et donc la production d'une mauvaise estimation du contour de sortie.

Nous avons également présenté l'estimation de données manquantes en utilisant l'information de la direction de la façade. Cette estimation aide à la détection finale de la ligne qui sépare la rue du trottoir derrière les obstacles. Nous avons seulement analysé la direction de la façade comme une seule ligne droite. Cependant, pour trouver la direction de bouts de rues, où il y a plusieurs directions de façades, d'autres modèles devront être étudiés. Finalement, la sortie de la segmentation de la chaussée est une image du contour comme l'empreinte au sol de la séparation. Une vectorisation du contour serait nécessaire pour permettre la modélisation de la chaussée

## **Conclusions et Perspectives**

---

### **Sommaire**

---

<b>13.1 Apports de cette thèse</b>	<b>172</b>
13.1.1 Ouverture ultime	172
13.1.2 Segmentation préalable à la modélisation urbaine	173
<b>13.2 Perspectives</b>	<b>174</b>
13.2.1 Validation	174
13.2.2 Chaîne de traitement	174
13.2.3 Combinaison des approches	175
13.2.4 Projet : Scan 3D en milieu urbain pour la mise à jour de plans	175

---

L'objectif principal de ce travail était d'étudier des approches morphologiques dans une étape de segmentation préalable à la modélisation d'environnements urbains. Tout au long de ce mémoire, nous avons présenté des techniques de segmentation d'images et des données laser afin d'alimenter une modélisation basée sur des règles. Nous rappellerons dans un premier temps les différents apports de cette thèse puis nous proposerons des directions de recherche dans la continuité de nos travaux.

### **13.1 Apports de cette thèse**

#### **13.1.1 Ouverture ultime**

Nous avons analysé de manière approfondie l'opérateur d'ouverture ultime en tant qu'opérateur de segmentation générique. Il s'agit d'un puissant outil non-paramétrique, intrinsèquement multi-échelle, permettant de préserver les structures les plus contrastées de l'image. Nous avons aussi identifié ses limitations dues aux structures imbriquées et aux transitions graduelles, en produisant des problèmes de masquage et de fuites des objets d'intérêt. Enfin nous avons analysé diverses pistes de réflexion concernant ses possibles extensions et nous en avons développé trois. Nous avons également proposé une mise en œuvre de ces stratégies, basée sur l'implémentation



de l'ouverture ultime en utilisant une structure *Max-Tree*, présenté dans [41].

La première extension définit un nouvel opérateur  $\theta^\Omega$  qui associe l'information de forme *a priori*, dans une application donnée, en utilisant une fonction de similarité. Dans la deuxième stratégie deux nouvelles définitions ont été proposées : l'ouverture ultime géodésique et l'ouverture ultime hiérarchique. Nous avons décrit l'ouverture ultime hiérarchique de manière générale ; néanmoins nous avons déterminé le niveau de hiérarchie optimale par rapport à un critère. D'ailleurs deux critères hiérarchiques ont été proposés : surface et hétérogénéité couleur. La dernière extension reprend la notion d'ouverture avec l'accumulation présentée dans [129]. Cet opérateur combat principalement le problème dû aux transitions graduelles de l'image, en assignant aux structures un contraste plus représentatif. Même si nous avons sacrifié le côté non paramétrique de l'opérateur classique, les trois opérateurs restent toutefois faciles à paramétrer. L'efficacité des trois approches a été démontrée dans le cadre de trois applications : segmentation d'images de façade, détection de texte enfoui et segmentation d'images de cellules. Les résultats de segmentation montrent de manière générale la diminution des problèmes de masquage et des fuites avec les trois stratégies proposées. Cette partie a donné lieu aux publications suivantes : [63, 64, 69].

### 13.1.2 Segmentation préalable à la modélisation urbaine

Nous avons proposé des méthodes d'analyse d'images et des données laser qui travaillent de manière indépendante. Concernant l'analyse d'images, l'utilisation de la projection du gradient n'est pas une méthode nouvelle, mais le fait d'utiliser plusieurs filtres morphologiques, telle que nous l'avons introduite dans nos travaux, constitue une manière originale d'utiliser la technique avec des façades haussmanniennes. De plus, la méthode permet d'analyser les images très texturées et même avec de petits obstacles. Cette technique de projection a été utilisée dans plusieurs étapes de l'analyse des images de façade. Nous avons d'abord proposé une méthode pour la séparation des façades qui composent une image. Ensuite des méthodes pour la décomposition d'une façade en étages et en travées sont aussi présentées. Finalement, l'analyse a été menée jusqu'à la détection de fenêtres et de balcons. Les résultats que nous obtenons sont très satisfaisants. Tous les outils que nous avons développés durant ce travail, ont bien répondu aux attentes qu'avaient les partenaires du projet. D'ailleurs ils ont été intégrés dans une plateforme de modélisation procédurale, pour la création de modèles de façade. Cette partie a donné lieu aux publications suivantes : [66, 70].

En ce qui concerne l'analyse de nuages de points, les méthodes s'appuient sur une image de profondeur et/ou une image d'accumulation, en utilisant une caméra "virtuelle". Nous avons présenté une méthode afin de séparer les nuages de points en îlots et puis en façades et sol. Si les méthodes présentées peuvent se révéler efficaces, les procédures implémentées dépendent encore de la précision et de la "propreté" des nuages de points. Nous faisons les hypothèses que les points qui appartiennent à la façade se trouvent principalement perpendiculaires au

plan  $XY$ . Si cela n'est pas le cas, des méthodes pour la correction automatique de plans de façades devront être étudiées.

Par ailleurs l'un des apports majeurs de nos travaux est la méthode de détection et de classification automatique des artefacts. Elle permet d'un côté le filtrage des artefacts afin de faciliter la segmentation de façades et de la chaussée, de l'autre côté la réintroduction des certains objets du mobilier urbain (lampadaires) afin d'incrémenter le réalisme visuel d'une scène modélisée. Bien que dans la littérature, des méthodes de détection d'objets sur des données Lidar aériennes existent, nous sommes vraiment dans une autre échelle d'analyse car nos données sont acquises au sol et avec des précisions d'environ de 20cm. Nous détectons de manière satisfaisante tous les artefacts visibles à partir d'un point de vue supérieur de la scène. Seuls les artefacts "cachés" (artefacts sous un arbre) ou en contact avec la façade ne sont pas encore bien détectés.

Dans cette partie, nous avons aussi développé une méthode de segmentation de la chaussée. La méthode détecte les variations de hauteur entre la route et le trottoir qui correspondent aux standards de la hauteur d'un trottoir.

Cette partie a donné lieu aux publications suivantes : [65, 67, 68].

Enfin, l'ensemble des outils proposés dans cette thèse sont complètement automatiques avec les mêmes paramètres, pour l'ensemble de données testées.

## 13.2 Perspectives

Cette thèse constitue le premier travail du Centre de Morphologie Mathématique dans le domaine de l'analyse de données comme étape préalable à la modélisation d'environnements urbains. Les résultats probants obtenus dans le cadre de cette thèse laissent espérer de nouveaux projets dans la continuité de ce travail.

### 13.2.1 Validation

Un des axes de progrès vise à la "validation" quantitative des performances et de la fiabilité des modules que nous avons présentés dans ce manuscrit. Les méthodes ont déjà été testées sur une base de données relativement grande, mais petite par rapport à la dimension d'une ville comme Paris. Par exemple, dans les méthodes d'analyse de nuages de points, d'autres types d'artefacts pourront être identifiés afin d'améliorer leur classification.

Par ailleurs, même si les méthodes d'extraction sémantique ont été développées afin de diviser des façades haussmanniennes, il serait intéressant d'évaluer leur performance avec un autre type d'architecture.

### 13.2.2 Chaîne de traitement

Le Chapitre 7 illustre les résultats de trois applications des extensions de l'ouverture ultime proposées dans cette thèse. Cependant, les résultats restent de simples segmentations qui pourront être utilisées par une chaîne de re-

connaissance. Ainsi, pour les trois applications, une vérité terrain est nécessaire. Cette vérité terrain permettrait à la fois de réaliser une étape d'apprentissage et d'établir une mesure d'évaluation de la segmentation. Dans le cas particulier de la détection de texte enfoui, nous proposons d'intégrer les nouveaux opérateurs dans la chaîne de traitement présentée dans [129].

### 13.2.3 Combinaison des approches

Une voie de recherche particulièrement intéressante à considérer est la combinaison des approches présentées. Pour l'instant, les trois parties (ouverture ultime, analyse d'image de façade et analyse de nuages de points) restent indépendantes. Nous pouvons seulement citer l'approche de segmentation des nuages de points en bâtiments qui exploite le résultat de segmentation obtenu sur l'image de façade.

Dans une première partie nous proposons de combiner les extensions de l'ouverture ultime entre elles. Par exemple, nous pouvons facilement imaginer l'ouverture ultime hiérarchique par  $\Delta$ -attributs, de cette manière le côté hiérarchique réduira le problème de masquage et le côté accumulation produira des structures plus contrastées.

Par ailleurs, il serait particulièrement intéressant de combiner les approches de l'ouverture ultime avec celles d'analyse d'images de façade dans le but de détecter les fenêtres de manière locale. Rappelons que la détection de fenêtres par projection du gradient est réalisée globalement dans l'image, tous les étages à la fois. Ainsi, en combinant ces deux méthodes nous pourrions détecter les composantes connexes des fenêtres localement. De cette manière, la nouvelle méthode combinerait la robustesse de l'approche globale et la précision de l'approche ouverture ultime.

Finalement, nous proposons d'étudier une segmentation multimodale des données images et laser. Il faut mentionner que pour réaliser ce type de segmentation, un recalage "parfait" des données est indispensable. La segmentation multimodale présenterait de meilleurs résultats lorsque nous segmenterions des éléments qui possèdent les deux types d'information, comme c'est le cas des balcons. D'ailleurs cette combinaison permettrait une localisation de fenêtres plus précise car le laser traverse les vitres et donc sur les nuages de points il y aurait une discontinuité importante à l'intérieure des fenêtres.

### 13.2.4 Projet : Scan 3D en milieu urbain pour la mise à jour de plans

Actuellement, un nouveau projet est en cours de montage. Celui-ci a pour but d'améliorer le processus de création et de mise à jour de cartes de voirie urbaine, en utilisant des méthodes de relevé mobile et d'analyse automatique de scènes. Ainsi, les méthodes de détection présentées dans cette thèse pourront être utilisées pour l'identification automatique des objets d'intérêt.

## Bibliographie

---

- [1] Registration and Integration of Textured 3-D Data. In *3DIM '97 : Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling* (Washington, DC, USA, 1997), IEEE Computer Society, p. 234.
- [2] ABUHADROUS, I. *Système Embarqué Temps Réel de Localisation et de Modélisation 3D par Fusion Multi-Capteurs*. PhD thesis, CAOR - École Mines Paris, Janvier 2005.
- [3] AKBARZADEH, A., FRAHM, J.-M., MORDOHAI, P., CLIPP, B., ENGELS, C., GALLUP, D., MERRELL, P., PHELPS, M., SINHA, S., TALTON, B., WANG, L., YANG, Q., STEWENIUS, H., YANG, R., WELCH, G., TOWLES, H., NISTER, D., AND POLLEFEYS, M. Towards Urban 3D Reconstruction from Video. *International Symposium on 3D Data Processing, Visualization and Transmission 0* (2006), 1–8.
- [4] ALEGRE, F., AND DELLAERT, F. A Probabilistic Approach to the Semantic Interpretation of Building Façades. Tech. rep., Georgia Institute of Technology, 2004.
- [5] ALI, H., SEIFERT, C., JINDAL, N., PALETTA, L., AND PAAR, G. Window Detection in Facades. In *14th International Conference on Image Analysis and Processing* (2007), pp. 837–842.
- [6] ALIAGA, D. G., ROSEN, P. A., AND BEKINS, D. R. Style Grammars for Interactive Visualization of Architecture. *IEEE Transactions on Visualization and Computer Graphics* 13, 4 (2007), 786–797.
- [7] BAILLARD, C. *Analyse d’Images Aériennes Stéréoscopiques pour la Restitution 3D des Milieux Urbains*. PhD thesis, École Nationale Supérieure des Télécommunications de Paris, octobre 1997.
- [8] BALTER, R. *Construction d’un Maillage 3D Évolutif et Scalable pour le Codage Vidéo*. PhD thesis, Université de Rennes 1, Mai 2005.
- [9] BECKER, S., AND HAALA, N. Combined Feature Extraction for Façade Reconstruction. In *ISPRS Workshop on Laser Scanning and SilviLaser, Espoo, Finland* (September 2007), pp. 44–50.
- [10] BENZECRI, J. P. Compression des Images Polychromes et Sensibilité au Contraste chromatique. *Les cahiers de l’analyse des données XVII*, 2 (1992), 241–254.
- [11] BESL, P. J., AND MCKAY, H. D. A Method for Registration of 3D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, 2 (1992), 239–256.

- 
- [12] BEUCHER, S. *Segmentation d'Images et Morphologie Mathématique*. PhD thesis, École des Mines de Paris, June 1990.
- [13] BEUCHER, S. Watershed, Hierarchical Segmentation and Waterfall Algorithm. In *Mathematical morphology and its applications to image processing*, J. Serra and P. Soille, Eds. Kluwer Academic Publishers, 1994, pp. 69–76.
- [14] BEUCHER, S. Numerical Residues. *Image and Vision Computing* 25, 4 (2007), 405–415.
- [15] BEUCHER, S., AND MEYER, F. Morphological Approach to Segmentation : Watershed Transformation. In *Mathematical Morphology in Image Processing*, E. R. Dougherty, Ed. Marcel Dekker, New York, 1993, ch. 12, pp. 433–481.
- [16] BIRCH, P. J., BROWNE, S. P., JENNINGS, V. J., DAY, A. M., AND ARNOLD, D. B. Rapid Procedural-Modelling of Architectural Structures. In *VAST '01 : Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage* (New York, NY, USA, 2001), ACM, pp. 187–196.
- [17] BORSOTTI, M., CAMPADELLI, P., AND SCHETTINI, R. Quantitative Evaluation of Color Image Segmentation Results. *Pattern Recognition Letters* 19, 8 (1998), 741–747.
- [18] BOULAASSAL, H., LANDES, T., GRUSSENMEYER, P., AND TARSHA-KURDI, F. Automatic Segmentation of Building Facades using Terrestrial Laser Data. *ISPRS Workshop on Laser Scanning and SilviLaser, Espoo, Finland* (May 2007), 65–70.
- [19] BRÉDIF, M., BOLDO, D., DESEILLIGNY, M. P., AND MAÎTRE, H. 3D Building Reconstruction with Parametric Roof Superstructures. In *IEEE International Conference on Image Processing* (2007), IEEE, pp. 537–540.
- [20] BREEN, E. J., AND JONES, R. Attribute Openings, Thinnings, and Granulometries. *Computer Vision and Image Understanding* 64, 3 (1996), 377–389.
- [21] BRETAR, F. *Couplage de Données Laser Aéroporté et Photogrammétriques pour l'Analyse de Scènes Tridimensionnelles*. PhD thesis, École Nationale Supérieure des Télécommunications de Paris, 2006.
- [22] BROSTOW, G. J., SHOTTON, J., FAUQUEUR, J., AND CIPOLLA, R. Segmentation and Recognition Using Structure from Motion Point Clouds. In *ECCV '08 : Proceedings of the 10th European Conference on Computer Vision* (Berlin, Heidelberg, 2008), Springer-Verlag, pp. 44–57.
- [23] BROWN, L. G. A Survey of Image Registration Techniques. *ACM Comput. Surv.* 24, 4 (1992), 325–376.
- [24] BRUN, X. *Numérisation et Modélisation 3D d'Environnements Routiers, Application pour la Détermination de la Visibilité Géométrique des Routes*. PhD thesis, CAOR - École Mines Paris, Dec 2007.
- [25] BRUN, X., DESCHAUD, J.-E., AND GOULETTE, F. On-the-way City Mobile Mapping Using Laser Range Scanner and Fisheye Camera. In *5th International Symposium on Mobile Mapping Technology (MMT' 07)*, Padova, Italy (May 2007).

- [26] CHARPIAT, G., FAUGERAS, O., AND KERIVEN, R. Approximations of Shape Metrics and Application to Shape Warping and Empirical Shape statistics. *Foundations of Computational Mathematics* 5, 1 (2005), 1–58.
- [27] CHUNG CHANG, C., AND JEN LIN, C. LIBSVM : A Library for Support Vector Machines, 2001.
- [28] CORNELIS, N., LEIBE, B., CORNELIS, K., AND GOOL, L. V. 3D City Modeling Using Cognitive Loops. In *International Symposium on 3D Data Processing, Visualization and Transmission* (Washington, DC, USA, 2006), IEEE Computer Society, pp. 9–16.
- [29] DAGORNE, A., OTTAVI, J.-Y., CASTEX, J.-M., AND JULIAN., M. Systèmes d’Informations Géographiques et Gestion du Territoire. 28 49, In actes du colloque SIG-GIS CARTAO, 1991.
- [30] ÁDAM RAKUSZ, LOVAS, T., AND ÁARPARD BARSÍ. Lidar-Based Vehicle Segmentation. Tech. rep., Department of Photogrammetry and Geoinformatics, Budapest University of Technology and Economics, 2004.
- [31] DEBEVEC, P. E. *Modeling and Rendering Architecture from Photographs*. PhD thesis, University of California at Berkeley, October 1996.
- [32] DEVEAU, M. *Modélisation 3D par Utilisation Conjointe de Données Laser et Image*. PhD thesis, Université René Descartes Paris 5 - MATIS, 2006.
- [33] DIAS, P., SEQUEIRA, V., VAZ, F., AND GONÇALVES, J. G. Registration and Fusion of Intensity and Range Data for 3D Modelling of Real World Scenes. *International Conference on 3D Digital Imaging and Modeling 0* (2003), 418.
- [34] DICK, A. R., TORR, P. H. S., AND CIPOLLA, R. Modelling and Interpretation of Architecture from Several Images. *International Journal of Computer Vision* 60, 2 (2004), 111–134.
- [35] DIMITRIADOU, E., HORNIK, K., LEISCH, F., MEYER, D., AND WEINGESSEL, A. *Misc Functions of the Department of Statistics (e1071 Package)*. TU Wien, February 2009.
- [36] DOLD, C., AND BRENNER, C. Automatic Matching of Terrestrial Scan Data as a Basis for the Generation of Detailed 3D City Models. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey* (July 2004).
- [37] DOLD, C., AND BRENNER, C. Registration of Terrestrial Laser Scanning Data using Planar Patches and Image Data. In *International Archives of Photogrammetry and Remote Sensing* (2006), pp. 25–27.
- [38] DONIKIAN, S. Les Systèmes d’Information Géographique et la 3D pour la Représentation de Données Urbaines. Tech. rep., Irisa/CNRS, 2002.
- [39] DUARTE, J. *Malagueira Grammar - Towards a Tool for Customizing Alvaro Siza’s Mass Houses at Malagueira*. PhD thesis, MIT School of Architecture and Planning, 2002.

- 
- [40] EBERT, D. S., MUSGRAVE, K. F., PEACHEY, D., PERLIN, K., AND WORLEY, S. *Texturing & Modeling : A Procedural Approach, Third Edition (The Morgan Kaufmann Series in Computer Graphics)*. Morgan Kaufmann, December 2002.
- [41] FABRIZIO, J., AND MARCOTEGUI, B. Fast Implementation of the Ultimate Opening. In *ISMM '09 : Proceedings of the nine International Symposium on Mathematical Morphology* (August 2009).
- [42] FISCHLER, M. A., AND BOLLES, R. C. Random Sample Consensus : A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24, 6 (1981), 381–395.
- [43] FISHER, R. A. The use of Multiple Measurements in Taxonomic Problems. *Annals Eugen.* 7 (1936), 179–188.
- [44] FOLEY, J. D., VAN DAM, A., FEINER, S. K., AND HUGHES, J. F. *Computer Graphics : Principles and Practice in C*, second ed. Addison-Wesley Professional, August 1995.
- [45] FÖRSTNER, W. 3D-City Models : Automatic and Semiautomatic Acquisition Methods. Tech. rep., 1999.
- [46] FRÜH, C., SAMMON, R., AND ZAKHOR, A. Automated Texture Mapping of 3D City Models with Oblique Aerial Imagery. In *International Symposium on 3D Data Processing, Visualization and Transmission* (Washington, DC, USA, 2004), IEEE Computer Society, pp. 396–403.
- [47] FRÜH, C., AND ZAKHOR, A. 3D Model Generation for Cities using Aerial Photographs and Ground Level Laser Scans. *IEEE Computer Vision and Pattern Recognition* 02 (2001), 31–38.
- [48] FRÜH, C., AND ZAKHOR, A. Constructing 3D City Models by Merging Aerial and Ground Views. *IEEE Computer Graphics Applications* 23, 6 (2003), 52–61.
- [49] FRÜH, C., AND ZAKHOR, A. Constructing 3D City Models by Merging Ground-Based and Airborne Views. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Los Alamitos, CA, USA, 2003), vol. 02, IEEE Computer Society, pp. 306–313.
- [50] FUCHS, F. *Contribution à la Reconstruction du Bâti en Milieu Urbain, à l'Aide d'Images Aériennes Stéréoscopiques à Grande Échelle. Etude d'une Approche Structurale*. PhD thesis, Université Rene Descartes - Paris V, avril 2001.
- [51] GALLAGHER, A. C., LUO, J., AND HAO, W. Improved Blue Sky Detection using Polynomial Model Fit. In *IEEE International Conference on Image Processing, ICIP'04* (2004), pp. 2367–2370.
- [52] GANSTER, B., AND KLEIN, R. An Integrated Framework for Procedural Modeling. In *Spring Conference on Computer Graphics* (Apr. 2007), M. Sbert, Ed., Comenius University, Bratislava, pp. 150–157.
- [53] GERTH, B., BERNDT, R., HAVEMANN, S., AND FELLNER, D. W. 3D Modeling for Non-Expert Users with the Castle Construction Kit v0.5. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage* (2005), pp. 1–9.

- [54] GLORY, E., MEAS-YEDID, V., PINSET, C., OLIVO-MARIN, J.-C., AND G.STAMON. Automatic Quality Evaluation of Color Image Segmentations in Cytology. In *15e congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle* (January 2006).
- [55] GOOL, L. V., ZENG, G., DEN BORRE, F. V., AND MÜLLER, P. Towards Mass-produced Building Models. In *PIA'07 : Photogrammetric Image Analysis* (September 2007), pp. 209–220.
- [56] GOULETTE, F., NASHASHIBI, F., AMMOUN, S., AND LAURGEAU, C. An Integrated on-Board Laser Range Sensing System for On-the-Way City and Road Modelling. *Revue française de photogrammétrie et de télédétection* 185 (2007), 78.
- [57] GRUEN, A., AND NEVATIA, R. Guest Eds., Special Issue on Automatic Building Extraction From Aerial Images. *Computer Vision and Image Understanding* 72, 2 (November 1998), 99–214.
- [58] GUIGUES, L. *Modèles Multi-Echelles pour la Segmentation d'Images*. PhD thesis, Université de Cergy-Pontoise, 2003.
- [59] HANBURY, A., AND SERRA, J. Mathematical Morphology in the HLS Colour Space. In *Proceedings of the British Machine Vision Conference 2001* (2001), BMVA, pp. 451–460.
- [60] HANSEN, J., AND JONAS, D. Airborne Laser Scanning or Aerial Photogrammetry for the Mine Surveyor. *AAM Surveys Inc.* (1999).
- [61] HARTLEY, R. I., AND ZISSERMAN, A. *Multiple View Geometry in Computer Vision*, second ed. Cambridge University Press, ISBN : 0521540518, 2004.
- [62] HAVEMANN, S. *Generative Mesh Modeling*. PhD thesis, Dept. of Computer Science, Braunschweig Tech. Univ., 2005.
- [63] HERNÁNDEZ, J., AND MARCOTEGUI, B. Segmentation of Façade Images using Ultimate Opening. In *31eme journee ISS France* (Fevrier 2008).
- [64] HERNÁNDEZ, J., AND MARCOTEGUI, B. Document Image Binarization using Ultimate attribute Opening. In *DIBCO 2009 - Document Image Binarization Contest in the framework of ICDAR*. July 2009.
- [65] HERNÁNDEZ, J., AND MARCOTEGUI, B. Filtering of Artifacts and Pavement Segmentation from Mobile LIDAR Data. In *ISPRS Workshop on Laser Scanning '09, Paris, France* (September 2009).
- [66] HERNÁNDEZ, J., AND MARCOTEGUI, B. Morphological Segmentation of Building Facade Images. In *IEEE International Conference on Image Processing* (November 2009).
- [67] HERNÁNDEZ, J., AND MARCOTEGUI, B. Point Cloud Segmentation towards Urban Ground Modeling. In *5th GRSS/ISPRS Joint workshop on remote sensing and data fusion over urban areas, Shanghai, China* (May 2009).



- [68] HERNÁNDEZ, J., AND MARCOTEGUI, B. Segmentation et Interprétation des Nuages de Points pour la Modélisation d'Environnements Urbain. In *Colloque Techniques laser pour l'étude des environnements naturels et urbains, Mans, France* (Janvier 2009).
- [69] HERNÁNDEZ, J., AND MARCOTEGUI, B. Ultimate Attribute Opening Segmentation with Shape Information. In *ISMM '09 : Proceedings of the nine International Symposium on Mathematical Morphology* (August 2009).
- [70] HERNÁNDEZ, J., MARCOTEGUI, B., AND MEYER, F. Morphological Facade Image Analysis for Building Modeling. In *32eme journee ISS France* (Fevrier 2009).
- [71] HINZ, S., BAUMGARTNER, A., AND EBNER, H. Modeling Contextual Knowledge for Controlling Road Extraction in Urban Areas. In *IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas* (2001), pp. 40–44.
- [72] HOHMANN, B., KRISPEL, U., HAVEMANN, S., AND FELLNER, D. CityFit - High-Quality Urban Reconstructions by Fitting Shape Grammars to Images and Derived Textured Point Clouds. In *Proceedings of the 3rd ISPRS International Workshop 3D-ARCH 2009 : "3D Virtual Reconstruction and Visualization of Complex Architectures" Trento, Italy* (February 2009).
- [73] HU, J., YOU, S., AND NEUMANN, U. Approaches to Large-Scale Urban Modeling. *IEEE Computer Graphics Applications* 23, 6 (2003), 62–69.
- [74] HU, J., YOU, S., AND NEUMANN, U. Integrating LiDAR, Aerial Image and Ground Images for Complete Urban Building Modeling. In *International Symposium on 3D Data Processing, Visualization and Transmission* (Washington, DC, USA, 2006), IEEE Computer Society, pp. 184–191.
- [75] HUBER, M., SCHICKLE, W., HIM, S., AND BAUMGARTNE, A. Fusion of LIDAR Data and Aerial Imagery for Automatic Reconstruction of Building Surfaces. In *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*. (May 2003).
- [76] IWANOWSKI, M., AND SOILLE, P. Morphological Refinement of an Image Segmentation. In *11th International Conference on Computer Analysis of Images and Patterns* (September 2005), pp. 538–545.
- [77] JANIL, R., DESMAISON, O., PEYRONNET, P., AND GUIARD, C. Caractérisation des Typologies et Styles de Façades Haussmannienne et Post-Haussmanniennes Parisiennes. Tech. rep., THALES SERVICES SAS. 1, rue du Général de Gaulle – Z.I. Les Beaux Soleils, Osny BP 226, 95523 Cergy Pontoise Cedex, 2009.
- [78] JOHANSSON, B. *Computer Vision using Rich Features - Geometry and Systems*. PhD thesis, Lund University, 2002.
- [79] JWA, Y., SOHN, G., TAO, V., AND CHO, W. An Implicit Geometric Regularization of 3D Building Shape Using Airborne LIDAR Data. In *International Society for Photogrammetry and Remote Sensing* (2008), pp. 69–76.

- 
- [80] KORAH, T., AND RASMUSSEN, C. Spatiotemporal Inpainting for Recovering Texture Maps of Occluded Building Facades. *IEEE Transactions on Image Processing* 16, 9 (September 2007), 2262–2271.
- [81] KORČ, F., AND FÖRSTNER, W. Interpreting Terrestrial Images of Urban Scenes using Discriminative Random Fields. In *Proc. of the 21st Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS)* (2008).
- [82] KORČ, F., AND FÖRSTNER, W. eTRIMS Image Database for Interpreting Images of Man-Made Scenes. Tech. Rep. TR-IGG-P-2009-01, April 2009.
- [83] LAI, K., AND FOX, D. 3D Laser Scan Classification using Web Data and Domain Adaptation. In *Proceedings of Robotics : Science and Systems* (Seattle, USA, June 2009).
- [84] LECHNER, T., WILENSKY, U., FELSEN, M., AND REN, P. Procedural Modeling of Urban Land Use in Cities. Tech. rep., Dept. Computer Science, Northwestern University, 2004.
- [85] LEE, S. C., JUNG, S. K., AND NEVATIA, R. Automatic Integration of Facade Textures into 3D Building Models with a projective geometry Based line clustering. 511–519.
- [86] LEE, S. C., JUNG, S. K., AND NEVATIA, R. Integrating Ground and Aerial Views for Urban Site Modeling. In *International Conference on Pattern Recognition* (Los Alamitos, CA, USA, 2002), vol. 4, IEEE Computer Society, pp. 107–112.
- [87] LEE, S. C., AND NEVATIA, R. Extraction and Integration of Window in a 3D Building Model from Ground View Images. *IEEE Computer Vision and Pattern Recognition* 02 (2004), 113–120.
- [88] LEITE, N., AND GUIMARAES, S. Morphological Residues and a General Framework for Image Filtering and Segmentation. *Journal on Applied Signal Processing* 4 (December 2001), 219–229.
- [89] LIN, C., AND NEVATIA, R. Building Detection and Description from a Single Intensity Image. *Computer Vision and Image Understanding* 72, 2 (1998), 101–121.
- [90] LINDBERG, T. Feature detection with automatic scale selection. *International Journal of Computer Vision* 30 (1998), 77–116.
- [91] LIU, H., WANG, Q., HUA, W., ZHOU, D., AND BAO, H. Building Chinese Ancient Architectures in Seconds. In *International Conference on Computational Science* (2) (2005), pp. 248–255.
- [92] LIU, J., AND YANG, Y. Multiresolution Color Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16, 7 (1994), 689–700.
- [93] LONCARIC, S. A Survey of Shape Analysis Techniques. *Pattern Recognition* 31 (1998), 983–1001.
- [94] LOWE, D. G. Object Recognition from Local Scale-Invariant Features. In *IEEE International Conference on Computer Vision* (Los Alamitos, CA, USA, 1999), vol. 2, IEEE Computer Society, pp. 1150–1157.
- [95] LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60 (2004), 91–110.

- [96] LUO, J., AND ETZ, S. P. A Physical Model-Based Approach to Detecting Sky in Photographic Images. *IEEE Transactions on Image Processing* 11, 3 (2002), 201–212.
- [97] MADHAVAN, R., AND HONG, T. Robust Detection and Recognition of Buildings in Urban Environments from LADAR Data. In *International Conference on Artificial Intelligence and Pattern Recognition* (Los Alamitos, CA, USA, 2004), IEEE Computer Society, pp. 39–44.
- [98] MAILLET, G., AND FLAMANC, D. Comparison of Aerial Images, Satellite Images and Laser Scanning DSM in a 3D City Models Production Framework. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. (July 2004), vol. 35.
- [99] MAINTZ, J. B., AND VIERGEVER, M. A. A Survey of Medical Image Registration. *Medical Image Analysis* 2, 1 (March 1998), 1–36.
- [100] MALLET, C., AND BRETAR, F. Full-waveform Topographic LiDAR : State-of-the-Art. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, 1 (January 2009), 1–16.
- [101] MARDIA, K. V., KENT, J. T., AND BIBBY, J. M. *Multivariate Analysis*. Academic Press, 1979.
- [102] MARVIE, J.-E., PERRET, J., AND BOUATOUCH, K. The FL-System : a Functional L-System for Procedural Geometric Modeling. *The Visual Computer* 21, 5 (June 2005), 329–339.
- [103] MAYER, H., AND REZNIK, S. Building Façade Interpretation from Image Sequences. In *CMRT05, Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms, and Evaluation* (2005), vol. XXXVI, pp. 55 – 60.
- [104] MAYER, H., AND REZNIK, S. MCMC Linked with Implicit Shape Models and Plane Sweeping for 3D Building Facade Interpretation in Image Sequences. In *Photogrammetric Computer Vision and Image Analysis* (2006).
- [105] MEIJSTER, A., AND WILKINSON, M. A Comparison of Algorithms for Connected Set Openings and Closings. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 4 (2002), 484–494.
- [106] MENDEZ, E., SCHALL, G., HAVEMANN, S., JUNGHANN, S., FELLNER, D., AND SCHMALSTIEG, D. Generating Semantic 3D Models of Underground Infrastructure. *IEEE Computer Graphics and Applications* 28, 3 (2008), 48–57.
- [107] MEYER, F. From Connected Operators to Levelings. In *ISMM '98 : Proceedings of the fourth International Symposium on Mathematical Morphology and its Applications to Image and Signal Processing* (Norwell, MA, USA, 1998), Kluwer Academic Publishers, pp. 191–198.
- [108] MEYER, F. Morphological Multiscale and Interactive Segmentation. In *NSIP'99 : Nonlinear Signal and Image Processing* (1999), pp. 369–377.
- [109] MIKOLAJCZYK, K., AND SCHMID, C. An Affine Invariant Interest Point Detector. In *International Journal of Computer Vision* (2002), pp. 128–142.

- [110] MIKOLAJCZYK, K., TUYTELAARS, T., SCHMID, C., ZISSERMAN, A., MATAS, J., SCHAFFALITZKY, F., KADIR, T., AND VAN GOOL, L. A Comparison of Affine Region Detectors. *International Journal of Computer Vision* 65, 1-2 (2005), 43–72.
- [111] MÜLLER, P. Applied Procedural Modeling. In *ACM SIGGRAPH 2007 courses* (New York, NY, USA, 2007), ACM, pp. 112–147.
- [112] MÜLLER, P., VEREENOOOGHE, T., WONKA, P., PAAP, I., AND GOOL, L. V. Procedural 3D Reconstruction of Puuc Buildings in Xkipché. In *VAST 2006 : 7th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage* (October 2006), pp. 139–146.
- [113] MÜLLER, P., WONKA, P., HAEGLER, S., ULMER, A., AND GOOL, L. V. Procedural Modeling of Buildings. *ACM Transaction on Graphics* 25, 3 (2006), 614–623.
- [114] MÜLLER, P., ZENG, G., WONKA, P., AND GOOL, L. V. Image-Based Procedural Modeling of Facades. *ACM Transactions on Graphics* 26, 3 (2007), 85–93.
- [115] NAGAO, M., MATSUYAMA, T., AND IKEDA, Y. Region Extraction and Shape Analysis in Aerial Photographs. *Computer Graphics and Image Processing* 10, 3 (July 1979), 195–223.
- [116] NAJMAN, L., AND COUPRIE, M. Building the Component Tree in Quasi-Linear Time. *IEEE Transactions on Image Processing* 15, 11 (2006), 3531–3539.
- [117] NAJMAN, L., AND SCHMITT, M. Geodesic Saliency of Watershed Contours and Hierarchical Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 12 (1996), 1163–1173.
- [118] OUTAL, S. *Quantification par Analyse d’Images de la Granulométrie des Roches Fragmentées : Amélioration de l’Extraction Morphologique des Surfaces, Amélioration a Reconstruction Stéréologique*. PhD thesis, CMM/GEOSCIENCES - École Mines Paris, 2006.
- [119] PAPARODITIS, N. *Reconstruction 3D de Paysages Péri Urbains en Imagerie Stéréoscopique Satellitale Haute Résolution*. PhD thesis, Université de Nice - Sophia Antipolis, janvier 1998.
- [120] PARISH, Y. I. H., AND MÜLLER, P. Procedural Modeling of Cities. In *SIGGRAPH ’01 : Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2001), ACM, pp. 301–308.
- [121] PATTERSON, A., MORDOHAJ, P., AND DANIILIDIS, K. Object Detection from Large-Scale 3D Datasets using Bottom-Up and Top-Down Descriptors. In *ECCV ’08 : Proceedings of the 10th European Conference on Computer Vision* (Marseille, France, October 2008), pp. IV : 553–566.
- [122] PERBET, F. Dynamic Graph : un outil générique pour la modélisation multi-échelle. In *16èmes Journées de l’Association Française d’Informatique Graphique, AFIG 2003, December, 2003* (Paris, France, Dec. 2003), pp. 141–150.

- [123] PERRET, J. *Modélisation d'Environnements Urbains Virtuels*. PhD thesis, Université de Rennes 1, campus de Beaulieu, 35042 Rennes cedex, nov 2006.
- [124] PIEGL, L. A., AND TILLER, W. *The Nurbs Book*. Springer, 1997.
- [125] POLLEFEYS, M., GOOL, L. V., VERGAUWEN, M., CORNELIS, K., VERBIEST, F., AND TOPS, J. 3D Recording for Archaeological Fieldwork. *IEEE Computer Graphics Applications* 23, 3 (2003), 20–27.
- [126] POLLEFEYS, M., KOCH, R., VERGAUWEN, M., AND GOOL, L. V. Hand-Held Acquisition of 3D Models with a Video Camera. In *IEEE proceedings of Second International Conference on 3-D Imaging and Modeling* (Los Alamitos, CA, USA, 1999), vol. 00, IEEE Computer Society, pp. 14–23.
- [127] POLLEFEYS, M., NISTER, D., FRAHM, J., AKBARZADEH, A., MORDOHAI, P., CLIPP, B., ENGELS, C., GALLUP, D., KIM, S., MERRELL, P., SALMI, C., SINHA, S., TALTON, B., WANG, L., YANG, Q., STEWENIUS, H., YANG, R., WELCH, G., AND TOWLES, H. Detailed Real-Time Urban 3D Reconstruction from Video. *International Journal of Computer Vision* 78, 2-3 (July 2008), 143–167.
- [128] PRUSINKIEWICZ, P., AND LINDENMAYER, A. *The Algorithmic Beauty of Plants (The Virtual Laboratory)*. Springer, October 1991.
- [129] RETORNAZ, T. *Détection de Textes Enfouis dans des Bases d'Images Généralistes. Un Descripteur Sémantique pour l'Indexation*. PhD thesis, École des Mines de Paris, 2007.
- [130] RETORNAZ, T., AND MARCOTEGUI, B. Scene Text Localization Based on the Ultimate Opening. In *ISMM '07 : Proceedings of the eighth International Symposium on Mathematical Morphology* (October 2007), vol. 1, pp. 177–188.
- [131] RICARD, J., ROYAN, J., AND AUBAULT, O. Visualization of Real Cities Based on Procedural Modeling. In *IEEE Virtual Reality Workshop on Virtual Cityscapes : Key Research Issues in Modeling Large-Scale Immersive Urban Environments* (2008).
- [132] RIPPERDA, N. Determination of Facade Attributes for Facade Reconstruction. In *Intern. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (2008), pp. 285–290.
- [133] RIPPERDA, N. Grammar Based Facade Reconstruction using rjcmc. In *Photogrammetrie Fernerkundung Geoinformation (PFG)* (2008), pp. 83–92.
- [134] RISSON, V. *Application de la Morphologie Mathématique à l'Analyse des Conditions d'Eclairage des Images Couleur*. PhD thesis, CMM - Mines ParisTech, 2001.
- [135] ROEVER, C., RAABE, N., LUEBKE, K., LIGGES, U., SZEPANNEK, G., AND ZENTGRAF, M. *Classification and Visualization (klaR Package)*. Fakultät Statistik, Technische Universität Dortmund, February 2009.
- [136] ROTTENSTEINER, F. Status and Further Prospects of Object Extraction from Image and Laser Data. In *5th GRSS/ISPRS Joint workshop on remote sensing and Data fusion over urban areas, Shanghai, China* (May 2009).

- [137] SALEMBIER, P. Morphological Multiscale Segmentation for Image Coding. In *Proceedings of of the IEEE winter workshop on Nonlinear digital signal processing* (Amsterdam, The Netherlands, The Netherlands, 1994), Elsevier Science Publishers B. V., pp. 359–386.
- [138] SALEMBIER, P., AND OLIVERAS, A. Practical Extensions of Connected Operator. In *ISMM '96 : Proceedings of the third International Symposium on Mathematical Morphology* (May 1996), p. 97–110.
- [139] SALEMBIER, P., OLIVERAS, A., AND GARRIDO, J. L. Anti-Extensive Connected Operators for Image and Sequence Processing. *IEEE Transactions on Image Processing* 7 (1998), 555–570.
- [140] SALEMBIER, P., AND SERRA, J. Flat Zones Filtering, Connected Operators, and Filters by Reconstruction. *IEEE Transactions on Image Processing* 4 (1995), 1153–1160.
- [141] SCHINDLER, G., AND DELLAERT, F. “4D Cities : Atlanta”. In 4D Cities Spatio-Temporal Reconstruction from Images. Tech. rep., Georgia Tech., May 2008.
- [142] SCHINDLER, G., DELLAERT, F., AND KANG, S. B. Inferring Temporal Order of Images from 3D Structure. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 0 (2007), 1–7.
- [143] SCHINDLER, K., AND BAUER, J. A Model-Based Method for Building reconstruction. In *HLK '03 : Proceedings of the First IEEE International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis* (Washington, DC, USA, 2003), IEEE Computer Society, p. 74.
- [144] SCHWALBE, E., MAAS, H., AND SEIDEL, F. 3D Building Model Generation from Airborne Laserscanner Data using 2D GIS Data and Orthogonal Point Cloud Projections. In *ISPRS Workshop on Laser Scanning* (2005).
- [145] SEQUEIRA, V., AND GONÇALVES, J. G. 3D Reality Modelling : Photo-Realistic 3D Models of Real World Scenes. In *International Symposium on 3D Data Processing Visualization and Transmission* (Los Alamitos, CA, USA, 2002), vol. 0, IEEE Computer Society, p. 776.
- [146] SERRA, J. Connectivity on Complete Lattices. *Journal of Mathematical Imaging and Vision* 9, 3 (1998), 231–251.
- [147] SERRA, J., AND VINCENT, L. An Overview of Morphological Filtering. *Circuits Syst. Signal Process.* 11, 1 (1992), 47–108.
- [148] SMITH, A. R. Plants, Fractals, and Formal Languages. *SIGGRAPH Comput. Graph.* 18, 3 (1984), 1–10.
- [149] SOCHMAN, J. Evaluation of the AdaBoost IPM. Tech. rep., Center for Machine Perception, Prague, Czech Republic, 2007.
- [150] SOILLE, P. *Morphological Image Analysis : Principles and Applications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.
- [151] SOILLE, P., AND GRATIN, C. An Efficient Algorithm for Drainage Networks Extraction on DEMs. *Journal of Visual Communication and Image Representation* 5, 2 (June 1994), 181–189.

- [152] SOILLE, P., AND GRAZZINI, J. Constrained Connectivity and Transition Regions. In *ISMM '09 : Proceedings of the 9th International Symposium on Mathematical Morphology and Its Application to Signal and Image Processing* (Berlin, Heidelberg, 2009), Springer-Verlag, pp. 59–69.
- [153] STAMOS, I. Automated 3D Modeling of Urban Environments. In *Proceedings of the 3rd ISPRS International Workshop 3D-ARCH 2009 : "3D Virtual Reconstruction and Visualization of Complex Architectures"* Trento, Italy (February 2009).
- [154] STAMOS, I., AND ALLEN, P. 3D Model Construction using Range and Image Data. In *IEEE Conference on Computer Vision and Pattern Recognition* (2000), vol. 1, pp. 531 – 536.
- [155] STAMOS, I., AND ALLEN, P. Automatic Registration of 2D with 3D Imagery in Urban Environments. In *International Conference on Computer Vision* (Los Alamitos, CA, USA, 2001), vol. 02, IEEE Computer Society, p. 731.
- [156] STAMOS, I., YU, G., WOLBERG, G., AND ZOKAI, S. 3D Modeling using Planar Segments and Mesh Elements. In *International Symposium on 3D Data Processing, Visualization and Transmission* (Washington, DC, USA, 2006), IEEE Computer Society, pp. 599–606.
- [157] STINY, G., AND GIPS, J. Shape Grammars and the Generative Specification of Painting and Sculpture. In *Information Processing* (Amsterdam, 1972), C. V. Friedman, Ed., pp. 1460–1465.
- [158] TAILLANDIER, F. *Reconstruction du Bâti en Milieu Urbain : une Approche Multi-vues*. PhD thesis, École Polytechnique, octobre 2004.
- [159] URBACH, E. R., BOERSMA, N. J., AND WILKINSON, M. H. Vector-Attribute Filters. In *ISMM '05 : Proceedings of the seven International Symposium on Mathematical Morphology* (Paris, 18-20 April 2005), pp. 95–104.
- [160] VAPNIK, V. N. *The Nature of Statistical Learning Theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [161] VELTKAMP, R. C., AND HAGEDOORN, M. State-of-the-Art in Shape Matching. Tech. rep., Principles of Visual Information Retrieval, 1999.
- [162] VELTKAMP, R. C., AND HAGEDOORN, M. Shape Similarity Measures, Properties, and Constructions. In *In Advances in Visual Information Systems, 4th International Conference, VISUAL 2000* (2000), Springer, pp. 467–476.
- [163] VERMA, V., KUMAR, R., AND HSU, S. 3D Building Detection and Modeling from Aerial LiDAR Data. In *CVPR '06 : Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Washington, DC, USA, 2006), IEEE Computer Society, pp. 2213–2220.
- [164] VINCENT, L. Morphological Area Openings and Closings for Grey-scale Images. In *Proceedings of the Workshop "Shape in Picture", 7-11 September 1992, Driebergen, The Netherlands* (Berlin, 1994), Y.-L. O, A. Toet, D. Foster, H. J. A. M. Heijmans, and P. Meer, Eds., pp. 197–208.

- [165] VINCENT, L., AND DOUGHERTY, E. R. Morphological Segmentation for Textures and Particles. In *Digital Image Processing Methods*, E. R. Dougherty, Ed. Marcel Dekker, New York, 1994, pp. 43–102.
- [166] VOSSELMAN, G. 3D Reconstruction of Roads and Trees for City Modelling. In *Proceedings of the ISPRS working group III/3 workshop 3D reconstruction from airborne laserscanner and InSAR data* (2003), G. V. H.-G. Maas and A. Streilein, Eds., vol. 34, Institute of Photogrammetry and Remote Sensing Dresden University of Technology, Dresden, Germany.
- [167] VOSSELMAN, G., AND DIJKMAN, S. 3D Building Model Reconstruction from Point Clouds and Ground Plans. *International Archives of Photogrammetry and Remote Sensing XXXIV-3/W4* (2001), 37–43.
- [168] VU, H., KERIVEN, R., LABATUT, P., AND PONS, J.-P. Towards High-resolution Large-Scale Multi-View Stereo. In *Conference on Computer Vision and Pattern Recognition* (Miami, June 2009).
- [169] WANG, X., AND HANSON, A. R. Surface Texture and Microstructure Extraction from Multiple Aerial Images. *Computer Vision and Image Understanding* 83, 1 (2001), 1–37.
- [170] WANG, X., TOTARO, S., TAILLANDIER, F., HANSON, A., AND TELLER, S. Recovering Façade Texture and Microstructure from Real-World Images. In *ECCV Texture 2002 Workshop, Copenhagen, Denmark* (2002).
- [171] WATSON, B., MÜLLER, P., VERYOVKA, O., FULLER, A., WONKA, P., AND SEXTON, C. Procedural Urban Modeling in Practice. *IEEE Computer Graphics Applications* 28, 3 (2008), 18–26.
- [172] WERNER, T., AND ZISSERMAN, A. New Techniques for Automated Architecture Reconstruction from Photographs. In *ECCV '02 : Proceedings of the 7th European Conference on Computer Vision-Part II* (2002), vol. 2, pp. 541–555.
- [173] WILKINSON, M. H. F., GAO, H., HESSELINK, W. H., JONKER, J.-E., AND MEIJSTER, A. Concurrent Computation of Attribute Filters on Shared Memory Parallel Machines. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 10 (2008), 1800–1813.
- [174] WILKINSON, M. H. F., AND ROERDINK, J. B. T. M. Fast Morphological Attribute Operations using Tarjan’s Union-Find Algorithm. In *ISMM '00 : Proceedings International Symposium on Mathematical Morphology* (2000), Kluwer, pp. 311–320.
- [175] YAO, W., HINZ, S., AND STILLA, U. Automatic vehicle extraction from airborne LiDAR Data of urban areas using Morphological reconstruction. In *IAPR Workshop on Pattern Recognition in Remote Sensing* (2008).
- [176] YIN, Z., ZHOU, X., BAKAL, C., LI, F., SUN, Y., PERRIMON, N., AND WONG, S. T. C. Using iterative cluster merging with improved gap statistics to perform online phenotype discovery in the context of high-throughput rna screens. *BMC Bioinformatics* 9 (2008), 264.



- [177] YU, G., GROSSBERG, M., WOLBERG, G., AND STAMOS, I. Think globally, cluster locally : A unified framework for range Segmentation. In *International Symposium on 3D Data Processing, Visualization and Transmission* (June 2008).
- [178] ZHANG, H., FRITTS, J. E., AND GOLDMAN, S. A. An Entropy-Based Objective evaluation Method for Image Segmentation. In *In Proceedings of the SPIE - Storage and Retrieval Methods and Applications for Multimedia* (2004), pp. 38–49.
- [179] ZHANG, Y. A Survey on Evaluation Methods for Image Segmentation. *Pattern Recognition* 29, 8 (August 1996), 1335–1346.
- [180] ZHANG, Z. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision* 13, 2 (1994), 119–152.
- [181] ZHAO, H., AND SHIBASAKI, R. Reconstructing textured CAD Model of urban environment using vehicle-borne laser range scanners and line cameras. In *ICVS '01 : Proceedings of the Second International Workshop on Computer Vision Systems* (London, UK, 2001), Springer-Verlag, pp. 284–297.

